



IFW

PATENT
81940.0070

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

MATSUNAMI, et al.

Serial No: 10/775,886

Filed: February 10, 2004

For: STORAGE DEVICE

Art Unit: 2655

Examiner: Not Assigned

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450, on
August 3, 2004

Date of Deposit

Shirley Ferguson

Name

Signature Date August 3, 2004

Signature

Date

TRANSMITTAL OF PRIORITY DOCUMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Enclosed herewith is a certified copy of Japanese patent application No. 2003-086828, which was filed March 27, 2003, from which priority is claimed under 35 U.S.C. § 119 and Rule 55.

Acknowledgment of the priority document(s) is respectfully requested to ensure that the subject information appears on the printed patent.

Respectfully submitted,

HOGAN & HARTSON L.L.P.

Date: August 3, 2004

By:

Lawrence J. McClure

Registration No. 44,228

Attorney for Applicant(s)

500 South Grand Avenue, Suite 1900
Los Angeles, California 90071
Telephone: 213-337-6700
Facsimile: 213-337-6701

340300194

ME

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出 願 年 月 日 2 0 0 3 年 3 月 2 7 日
Date of Application:

出 願 番 号 特 願 2 0 0 3 - 0 . 8 6 8 2 8
Application Number:

[ST. 10/C] : [J P 2 0 0 3 - 0 8 6 8 2 8]

願 人 株式会社日立製作所
Applicant(s):

BEST AVAILABLE COPY

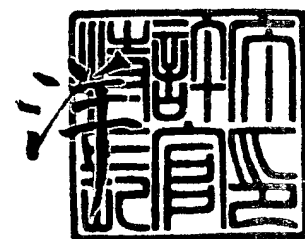
CERTIFIED COPY OF
PRIORITY DOCUMENT

CERTIFIED COPY OF
PRIORITY DOCUMENT

2 0 0 4 年 7 月 1 4 日

特許庁長官
Commissioner,
Japan Patent Office

小 川



【書類名】 特許願
【整理番号】 K03001941A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 12/00

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

【氏名】 松並 直人

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

【氏名】 藺田 浩二

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

【氏名】 山本 彰

【発明者】

【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所ストレージ事業部内

【氏名】 野沢 正史

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所システム開発研究所内

【氏名】 岩寄 正明

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶装置

【特許請求の範囲】

【請求項 1】

一又は複数の計算機と接続される記憶装置であって、
計算機からファイルの識別情報を指定したアクセス要求を受信する第一のインタフェース制御装置と、

前記第一のインタフェース制御装置と接続される第二のインタフェース制御装置と、

前記第二のインタフェース制御装置に接続される複数のディスクとを有し、

前記複数のディスクには、一又は複数の第一のディスクと、一又は複数の第二のディスクが含まれており、第一のディスクと第二のディスクは種類の異なるディスクであり、

前記第一のインタフェース制御装置は、計算機から受信した識別情報に基づいて、該識別情報によって指定されるファイルのデータの、前記複数のディスク内での格納位置を決定し、

前記第二のインタフェース制御装置は、前記第一のインタフェース制御装置によって決定された格納位置に、前記識別情報によって指定されるファイルのデータを格納するよう制御することを特徴とする記憶装置。

【請求項 2】

請求項 1 記載の記憶装置において、

前記第一のディスクはファイバチャネル型のインタフェースを備えるファイバチャネルディスクであり、前記第二のディスクはシリアル ATA 型のインタフェースを備えるシリアル ATA ディスクであることを特徴とする記憶装置。

【請求項 3】

請求項 1 記載の記憶装置において、

更に、メモリと、

前記メモリを制御するためのメモリコントローラと、

各々前記メモリコントローラに接続される、複数の第一のインタフェース制御

装置と、

各々前記メモリコントローラに接続される、複数の第二のインタフェース制御装置とを有し、

計算機からファイルの識別情報とファイルのデータを受信した第一のインタフェース制御装置は、ファイルのデータを前記メモリに格納し、

前記第一のインタフェース制御装置によって決定される前記ファイルのデータの格納位置に応じて、前記ファイルのデータが格納されるディスクに接続されている第二のインタフェース制御装置が、前記メモリに格納されている前記ファイルのデータをディスクに格納するよう制御することを特徴とする、記憶装置。

【請求項 4】

請求項 1 記載の記憶装置において、

前記一又は複数の第一のディスクには第一の記憶領域が存在し、

前記一又は複数の第二のディスクには第二の記憶領域が存在し、

前記第一のインタフェース制御装置は、前記第一の記憶領域に第一のファイルシステムを、前記第二の記憶領域に第二のファイルシステムを設定することを特徴とする記憶装置。

【請求項 5】

請求項 4 記載の記憶装置において、

前記第一のインタフェース制御装置は、計算機から受信した識別情報によって指定されるファイルの、予め定められている特徴である静的特性と、該ファイルが作成された時点からの時間の経過と共に変化する特徴である動的特性とに応じて、前記第一の記憶領域若しくは前記第二の記憶領域のいずれに、前記識別情報が示すファイルのデータを格納するかを決定することを特徴とする記憶装置。

【請求項 6】

請求項 5 記載の記憶装置において、

前記第一のインタフェース制御装置は、前記動的特性の変化に応じて、前記第一の記憶領域若しくは前記第二の記憶領域のいずれか一方に格納されているファイルのデータを、他方の記憶領域に移行するよう制御し、

該ファイルを特定するための識別情報と、該ファイルの格納位置との対応関係

を示す情報を変更することを特徴とする記憶装置。

【請求項 7】

請求項 6 記載の記憶装置において、

前記静的特性には、ファイルの種類を特定するための情報、若しくは、ファイルが作成された日付を特定するための情報、若しくは、ファイルの価値を特定するための情報が含まれており、

前記動的特性には、ファイルへのアクセス特性に関する情報、若しくは、ファイルが作成されてからの経過時間に関する情報が含まれていることを特徴とする記憶装置。

【請求項 8】

計算機に接続される記憶装置であって、

前記計算機からファイルの識別情報を有するアクセス要求を受信する一又は複数の第一のインタフェース制御装置と、

前記一又は複数の第一のインタフェース制御装置に接続される一又は複数の第二のインタフェース制御装置と、

各々が、前記一又は複数の第二のインタフェース制御装置のいずれかに接続される、複数の第一のディスクとを有し、

前記一又は複数の第一のインタフェース制御装置のいずれかは、複数の第二のディスクを有する第二の記憶装置に接続され、

前記複数の第一のディスクには第一の記憶領域が設定されており、

前記複数の第二のディスクには第二の記憶領域が設定されており、

前記計算機からアクセス要求を受信した第一のインタフェース制御装置は、受信したアクセス要求に含まれる識別情報によって指定されるファイルの特性に応じて、該ファイルのデータを前記第一の記憶領域若しくは前記第二の記憶領域のいずれに格納するかを決定し、

前記ファイルのデータが前記第一の記憶領域に格納される場合には、前記一又は複数の第二のインタフェース制御装置のいずれかが、該ファイルのデータを前記複数の第一のディスクのいずれかに格納し、

前記ファイルのデータが前記第二の記憶領域に格納される場合には、前記第二

の計算機と接続されている第一のインタフェース制御装置を介して、前記ファイルのデータを前記第二の記憶装置に送信するよう、前記計算機からアクセス要求を受信した第一のインタフェース制御装置が制御することを特徴とする記憶装置。

【請求項 9】

請求項 8 記載の記憶装置において、

前記第二の記憶装置は、ファイルの識別情報を有するアクセス要求を受信して、受信した識別情報に対応付けられた前記第二の記憶領域内の記憶領域にアクセスすることによって、該識別情報によって特定されるファイルのデータにアクセスする、第三のインタフェース制御装置を有し、

前記第三のインタフェース制御装置は、前記第二の記憶領域にファイルシステムを設定しており、

前記計算機からアクセス要求を受信した第一のインタフェース制御装置は、該アクセス要求に含まれる識別情報によって指定されるファイルのデータを、前記第二の記憶領域に格納する場合に、前記第二の記憶装置に接続されている第一のインタフェース制御装置を介して、該ファイルに対応付けられている識別情報を有するアクセス要求を、前記第三のインタフェース制御装置に送信するよう制御することを特徴とする記憶装置。

【請求項 10】

請求項 9 記載の記憶装置において、

前記第二の記憶装置と接続されている第一のインタフェース制御装置が、前記計算機からアクセス要求を受信することを特徴とする記憶装置。

【請求項 11】

請求項 9 記載の記憶装置において、

第一のインタフェース制御装置は、前記第一の記憶領域にデータが格納されているファイルの特性に基づいて、該ファイルのデータを前記第一の記憶領域から前記第二の記憶領域に、前記第三のインタフェース制御装置を介して移行するよう制御することを特徴とする記憶装置。

【請求項 12】

請求項 1 1 記載の記憶装置において、

第一のインタフェース制御装置は、前記第一の記憶領域に格納されているファイルのデータを、前記第二の記憶領域に移行する場合に、該ファイルに対応付けられている識別情報を有するアクセス要求を、前記第三のインタフェース制御装置に送信し、前記計算機から受信する該ファイルの識別情報と前記第二の記憶領域に設定されているファイルシステムとを対応付けて記憶することを特徴とする記憶装置。

【請求項 1 3】

請求項 1 2 記載の記憶装置において、

第一のインタフェース制御装置は、ファイルの特性に関する情報と前記第一の記憶領域及び前記第二の記憶領域の特性に関する情報を記憶しており、ファイルの特性に関する情報と、前記第一の記憶領域及び前記第二の記憶領域の特性に関する情報とに基づいて、前記第一の記憶領域に格納されているファイルのデータを前記第二の記憶領域に移行するか否かを決定することを特徴とする記憶装置。

【請求項 1 4】

計算機に接続される記憶装置であって、

前記計算機からファイルを指定するための識別情報を有するアクセス要求を受信する第一のインタフェース制御装置と、

前記第一のインタフェース制御装置に接続される第二のインタフェース制御装置と、

前記第二のインタフェース制御装置に接続される複数の第一のディスクと、

データの格納位置を示すアドレス情報が含まれるアクセス要求を受信する第四のインタフェース制御装置と、該第四のインタフェース制御装置に接続される複数の第二のディスクとを有する第二の記憶装置に接続される、第三のインタフェース制御装置とを有し、

前記複数の第一のディスクには第一の記憶領域が存在し、

前記複数の第二のディスクには第二の記憶領域が存在し、

前記計算機からアクセス要求を受信した第一のインタフェース制御装置は、受信したアクセス要求に含まれる識別情報が示すファイルの特性に応じて、該ファ

イルのデータを前記第一の記憶領域若しくは前記第二の記憶領域のいずれに格納するかを決定し、

前記ファイルのデータが前記第一の記憶領域に格納される場合には、前記第二のインタフェース制御装置が、該ファイルのデータを前記複数の第一のディスクのいずれかに格納し、

前記ファイルのデータが前記第二の記憶領域に格納される場合には、前記第一のインタフェース制御装置が、該ファイルのデータが格納される前記第二の記憶領域内のアドレス情報を含むアクセス要求を、前記第三のインタフェース制御装置を介して、前記第四のインタフェース制御装置に送信するよう制御することを特徴とする記憶装置。

【請求項 1 5】

請求項 1 4 記載の記憶装置において、

前記第三のインタフェース制御装置は、ブロック I/O インタフェースに対応するインタフェース制御装置であることを特徴とする記憶装置。

【請求項 1 6】

請求項 1 5 記載の記憶装置において、

前記第一のインタフェース制御装置は、前記第二の記憶領域に、ファイルシステムを設定することを特徴とする記憶装置。

【請求項 1 7】

請求項 1 6 記載の記憶装置において、

前記第一のインタフェース制御装置は、前記第一の記憶領域にデータが格納されているファイルの特性に基づいて、該ファイルのデータを前記第一の記憶領域から前記第二の記憶領域に、前記第三のインタフェース制御装置を介して移行するよう制御することを特徴とする記憶装置。

【請求項 1 8】

請求項 1 7 記載の記憶装置において、

前記第一のインタフェース制御装置は、前記第一の記憶領域に格納されているファイルのデータを、前記第二の記憶領域に移行する場合に、該ファイルのデータが格納される前記第二の記憶領域内の記憶領域のアドレスを含むアクセス要求

を、前記第三のインタフェース制御装置を介して前記第二の記憶装置に送信するよう制御し、

該ファイルの識別情報と、該ファイルのデータが格納される記憶領域を示す情報との対応関係を変更することを特徴とする記憶装置。

【請求項 19】

計算機に接続される記憶装置において、

前記計算機からファイルの識別情報を有するアクセス要求を受け付ける第一のノードと、

一又は複数の第一のディスクに接続される第二のノードと、

一又は複数の第二のディスクと、該一又は複数の第二のディスクに接続され、ファイルの識別情報を有するアクセス要求を受け付けるファイルI/Oインタフェース制御装置を有する第二の記憶装置に接続される、第三のノードと、

一又は複数の第三のディスクと、該一又は複数の第三のディスクに接続され、データの該一又は複数の第三のディスク内での格納位置を示すアドレス情報を有するアクセス要求を受け付けるブロックI/Oインタフェース制御装置を有する第三の記憶装置に接続される、第四のノードと、

前記第一のノードと、前記第二のノードと、前記第三のノードと、前記第四のノードとを相互に接続するスイッチとを有し、

前記一又は複数の第一のディスク内には第一の記憶領域が、前記一又は複数の第二のディスク内には第二の記憶領域が、前記一又は複数の第三のディスク内には第三の記憶領域が存在し、

前記第一のノードは、前記計算機から受信した識別情報によって特定されるファイルの特性に応じて、該ファイルを前記第一の記憶領域、前記第二の記憶領域、前記第三の記憶領域のいずれかに格納するよう制御することを特徴とする記憶装置。

【請求項 20】

請求項 19 記載の記憶装置において、

ファイルのデータを前記第一の記憶領域に格納する場合には、前記計算機から受信した識別情報に対応付けられている、前記一又は複数の第一のディスク内の

記憶領域に、前記第二のノードが、該ファイルのデータを格納するよう制御し、

ファイルのデータを前記第二の記憶領域に格納する場合には、前記第一のノードが、該ファイルに対応付けられた識別情報を含むアクセス要求を、前記第三のノードを介して前記第二の記憶装置に送信するよう制御し、

ファイルのデータを前記第三の記憶領域に格納する場合には、前記第一のノードが、該ファイルのデータの前記第三の記憶領域内での格納位置を示すアドレス情報を含むアクセス要求を、前記第四のノードを介して前記第三の記憶装置に送信するよう制御することを特徴とする記憶装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、計算機システムで用いられる記憶装置に関する。

【0002】

【従来の技術】

特許文献1には、階層記憶装置とよぶ計算機に、高速記憶装置と低速記憶装置を接続したシステムが開示されている。特許文献1では、使用頻度の高いファイルは磁気ディスク装置などの高速記憶装置に格納され、使用頻度の低いファイルはテープ装置などの安価な低速記憶装置に格納される。そして、ファイル毎のアクセス頻度を管理するテーブルを用いて、どのファイルをどの記憶装置に配置する、即ち格納するかが決定されている。

【0003】

また、特許文献2では、ディスクアレイサブシステムに代表されるような、計算機に接続して使用される記憶装置の内部に、処理速度や記憶容量の異なる複数の論理記憶装置を構成したシステムが開示されている。記憶装置は、記憶装置に格納されているデータに対する計算機からのアクセス頻度を、統計情報として管理し、この統計情報を基に、アクセス頻度の高いデータをより高性能な論理記憶装置へ移動させるための技術を開示している。

【0004】

【特許文献1】

特開平 9 - 2 9 7 6 9 9 号公報（第 3 - 4 頁、図 1）

【特許文献 2】

特開平 0 9 - 2 7 4 5 4 4 号公報（第 3 頁、図 6、7）

【0 0 0 5】

【発明が解決しようとする課題】

従来の技術における第一の課題は、記憶装置に接続される計算機への依存度が高く、システム構成に制約があり、システム管理の容易化が困難であることである。

【0 0 0 6】

特許文献 1 では、計算機上で動作するソフトウェアにより階層記憶制御を実現している。ここで階層記憶制御とは、処理速度や記憶容量の異なる複数の記憶領域に、データの使用頻度に応じて記憶領域を変更できるよう制御するデータ記憶制御のことである。即ち、階層記憶制御とは、データの使用頻度等データの特性に応じて、処理速度や記憶容量等の特性が異なる複数の記憶領域から適切な記憶領域を選択して、選択された記憶領域にデータを記憶させるよう制御することである。しかし古い計算機を新しい計算機にリプレイスする場合等、システム構成を変更しようとした際に、ソフトウェアの制御情報を引き継ぐことができない等の理由により、システムの維持が困難になることがある。

【0 0 0 7】

特許文献 2 では、論理記憶装置単位で階層記憶制御が実施されているが、記憶装置が論理記憶装置に記憶されているデータのデータ構造を認識するための技術や、排他制御を実行するための技術は開示されていない。従って、複数の計算機で同一の論理記憶装置を共有することは困難であり、複数の計算機によって使用される記憶装置を集約化して、計算機システムの管理コストを低減するためには、計算機毎に論理記憶装置を割り当てる等、計算機システムの構成に一定の制約が必要となる。

【0 0 0 8】

第二の課題は、データのライフサイクルや種類に応じたデータの最適配置が困難であることである。

【0 0 0 9】

従来技術では、過去にアクセス頻度が高かったデータは、将来もアクセス頻度が高いものと推定し、データのアクセス頻度に関する統計情報や、高速にアクセスできる記憶領域の使用容量に応じて、データを記憶させる記憶領域を決定している。そして、アクセス頻度の高いデータが高速にアクセスできる記憶装置に存在する確率を高めることで、処理効率の向上を図っている。しかし、データのライフサイクル、すなわちファイル生成後からの経過時間や、データを生成したり使用したりするアプリケーションの種類やデータ自体の種類に依存する、データの特性の相違を考慮して、データを記憶させる記憶領域を決定する技術は開示されていない。

【0 0 1 0】

第三の課題は、階層記憶制御の効果が小さいことである。

【0 0 1 1】

特許文献 1 では、磁気テープと磁気ディスクの容量差と価格差を利用して階層記憶制御を実行しているが、近年、磁気テープと磁気ディスクの容量差、価格差は縮小しており、これに伴って階層記憶制御によるコストの適正化、低コスト化の効果は小さくなっている。更に、磁気テープへのアクセススピードは磁気ディスクへのアクセススピードに比べて非常に遅いので、磁気テープをオンラインアクセス用の記憶装置に用いるのは困難である。

特許文献 2 では、磁気ディスクの R A I D 構成の違いによる価格差と性能差を利用して階層記憶制御を実行しているが、価格差は RAID 構成における冗長さの差に起因するだけなので、冗長さの差分のみの低コスト化しか望めない。

【0 0 1 2】

本発明の目的は、ホスト計算機で実行される O S やアプリケーションに依存することなく、ファイルの格納位置についての階層記憶制御が実行できる制御方法又は記憶装置を提供することである。

【0 0 1 3】

本発明の他の目的は、複数の計算機からファイルを共用することができる階層記憶制御方法、又は階層記憶制御を実行する記憶装置を提供することである。

【0014】

本発明の他の目的は、ファイルの特性に応じた階層記憶制御を実行できる制御方法又は記憶装置を提供することである。

【0015】

本発明の他の目的は、低コスト化の効果の高い階層記憶制御方法、又は階層記憶制御を実行する記憶装置を提供することである。

【0016】**【課題を解決するための手段】**

記憶装置は、特性の異なる複数の記憶領域と、一又は複数の計算機からファイルの識別情報を有するアクセス要求を受け付けるインタフェース制御装置と、識別情報によって指定されるファイルのデータが格納されている記憶領域にアクセスするためのインタフェース制御装置とを有し、インタフェース制御装置がファイルの特性に応じて、ファイルのデータを複数の記憶領域のいずれかに格納するよう制御する。

【0017】**【発明の実施の形態】**

本発明の実施形態を以下に説明する。尚本実施形態によって本発明が限定されるものではない。

[実施形態1]**(1) システム構成の一例 (図1)**

図1は、本発明が適用される、記憶装置1を含む計算機システムの一例を示した図である。以下、 x は任意の整数を表す。

【0018】

記憶装置1は、ディスクコントローラ（以下、「DKC」と称する）11及び複数の磁気ディスク装置（以下単にディスクと称する）170 x 、171 x を有するディスクアレイシステムである。本実施形態においては、記憶装置1は2種類のディスク170 x 、171 x を備えている。170 x はFibre Channel（以下、「FC」と称する）型のインタフェースを備えるFCディスク、171 x はserial AT Attached（以下、「SATA」と称する）型のインタフェースを備えるSATAディスクである。そして、複数のFC

ディスク170xはFCディスクプール 0 (170)を、複数のSATAディスク171xはSATAディスクプール 1 (171)を構成している。ディスクプールの詳細は後述する。

【 0 0 1 9 】

次に、記憶装置 1 のDKC11の構成を説明する。DKC11は、一又は複数のNASチャネルアダプタ110x、一又は複数のファイバチャネルアダプタ111x、複数のディスクアダプタ12x、共有メモリ13（以下「SM」と呼ぶ。）、共有メモリコントローラ15（以下「SMC」と呼ぶ。）、キャッシュメモリ14（以下「CM」と呼ぶ。）、キャッシュメモリコントローラ16（以下「CMC」と呼ぶ）を有する。

【 0 0 2 0 】

NASチャネルアダプタ（以下「CHN」と称する）110xは、ローカルエリアネットワーク（以下LANと称する。）20やLAN21と接続される計算機40x（以下、NASホストと称する）と、ファイル I / O インタフェースで接続されるインタフェース制御装置である。

【 0 0 2 1 】

ファイバチャネルアダプタ（以下、「CHF」と称する）111xは、ストレージエリアネットワーク（以下SANと称する。）30に接続される計算機（以下、SANホストと称する）50xと、ブロック I / O インタフェースで接続されるインタフェース制御装置である。以下、CHN及びCHFをまとめてチャネルアダプタ（以下「CH」と称する。）と称する。

【 0 0 2 2 】

ディスク17xは、ディスクアダプタ12xに接続されている。各ディスクアダプタ（以下、「DKA」と称する）12xは、自身に接続されている一又は複数のディスク17xへの入出力を制御する。

【 0 0 2 3 】

SMC15は、CHN110x、CHF111x、DKA12x及びSM13と接続される。SMC15は、CHN110x、CHF111x、及びDKA12xとSM13との間のデータ転送を制御する。CMC16は、CHN110x、CHF111x、DKA12x及びCM14と接続される。CMC16は、CHN110x、CHF111x、及びDKA12xとCM14との間のデータ転送を制御する。

【 0 0 2 4 】

SM13は、ディスクプール管理テーブル131を記憶している。ディスクプール管理テーブル131は、ディスクプールの構成を管理するために用いられる情報である。

【0025】

LAN20及び21は、CHN110xとNASホスト40xとを接続する。一般的に、LANにはEthernet（登録商標）が用いられる。SAN30は、CHF111x及びSANホスト50xとを接続する。一般的に、SANにはFibre Channelが用いられる。しかしSANとしてIPネットワークを用い、SANに接続される機器間ではSCSIプロトコルに従ったSCSIコマンドをIPパケットでカプセル化して送受信する、iSCSIを用いることとしても良い。SAN35は、本実施形態においては、記憶装置を接続するための専用のSANであるものとし、SAN35にはSANホストは接続していないものとする。

【0026】

記憶装置1では、すべてのCHが、CMC16又はSMC15を介して、CM14、SM13、全てのDKA12x、及びすべてのディスク17xへアクセスすることが出来る。

【0027】

尚図1に示す記憶装置1は、SANホスト50xに接続するためのSANインタフェース(CHF111x)と、NASホスト40xと接続するためNASインタフェース(CHN110x)の双方を有するが、本実施形態は、記憶装置1がNASインタフェースのみを有する場合であっても、実施可能な実施形態である。

【0028】

(2) 記憶装置の外観例 (図2)

図2は、記憶装置1の外観の一例を示す図である。

【0029】

DKCユニット19は、DKC11の構成要素である、CHN110x、CHF111x、DKA12x、SM13、及びCM14を格納する。SM13は、実際には複数のコントローラボード13xで構成される。また、CM14も、複数のキャッシュボード14xで構成される。記憶装置1の使用者は、これらのボードの枚数を増減して、所望の記憶容量のCM14又はSM13を有する記憶装置1を構成することができる。ディスクユニット(以下、「DKU」と称する)180及びDKU181は、ディスクプール170及びディスクプール171を格納す

る。

【0030】

スロット190には、CHN110x、CHF111x、DKA12x、コントローラボード13x、キャッシュボード14x等が作りこまれたアダプタボードが格納される。本実施形態においては、スロット190の形状、アダプタボードのサイズ及びコネクタの形状を、アダプタボードの種類やインタフェースの種類を問わず一定にし、互換性を保つようにする。したがって、DKCユニット19には、アダプタボードの種類、インタフェースの種類を問わず、任意のスロット190に任意のアダプタボードを装填することができる。また、記憶装置1の利用者は、CHN110x及びCHF111xのアダプタボードの数を自由に選択して、DKCユニット19のスロットに選択された数のCHN及びCHFを装填することが出来る。

【0031】

(3)CHN110xが作りこまれたアダプタボード（以下「NASボード」と称する。）の外観構成の一例（図3）

図3は、NASボードの外観構成の一例を示す図である。コネクタ11007は、DKCユニット19が有するコネクタと接続される。インタフェースコネクタ2001は、イーサネット（登録商標）に対応し、イーサネット（登録商標）と接続可能である。

【0032】

本実施形態においては、上述したように、アダプタボードのコネクタの形状は、アダプタボードの種類によらず一定であるため、CHN110xが作りこまれたアダプタボードとCHF111xが作りこまれたアダプタボードとは同一形状のコネクタを有する。尚、CHF111xが作りこまれたアダプタボードの場合、インタフェースコネクタ2001は、ファイバチャネルに対応し、ファイバチャネルと接続できるように構成されている。

【0033】

(4)NASボード(若しくはCHN)の構成の一例（図4）

図4は、CHN110xの構成の一例を示す図である。ファイルアクセス制御用CPU11001は、ファイルアクセスを制御するプロセッサである。LANコントローラ11002

は、インタフェースコネクタ2001を介してLANと接続され、LANとの間のデータの送受信を制御する。ファイルアクセス制御用メモリ11004は、ファイルアクセス制御用CPU11001と接続される。ファイルアクセス制御用メモリ11004には、ファイルアクセス制御用CPU11001が実行するプログラムや、制御データが格納されている。

【0034】

ディスクアレイ制御用CPU11008は、ディスクアレイを制御するプロセッサである。ここで、ディスクアレイとは、複数のディスクから構成される記憶装置のことである。特に複数のディスクのうち少なくとも1台以上に冗長データを格納し、耐障害性を備えたディスクアレイのことをRAIDと呼ぶ。RAIDについては後述する。ディスクアレイ制御用メモリ11009は、ディスクアレイ制御用CPU11008と接続され、ディスクアレイ制御用CPU11009が実行するプログラムや、制御データが格納されている。SM I/F制御回路11005は、CHN110xからSM13へのアクセスを制御する回路である。CM I/F制御回路11006は、CHN110xからCM14へのアクセスを制御する回路である。CPU間通信回路11007は、ファイルアクセス制御用CPU11001がディスクアクセスのためにディスクアレイ制御用CPU11008と通信する際に用いられる通信回路である。

【0035】

なお、本実施形態では、CHNにファイルアクセス制御用CPU11001とディスクアレイ制御用CPU11008の2つのプロセッサが搭載された、非対象型のマルチプロセッサ構成の例を示しているが、ファイルアクセス制御とディスクアレイ制御を単一のプロセッサで実行するようCHNを構成して実装することもできるし、ファイルアクセス制御とディスクアレイ制御を二つ以上のプロセッサで対等に実行する対称型マルチプロセッサ構成としてCHNを実装することもできる。

【0036】

尚、CHF111xの構成は、図4の上半分に示された構成要素である、LANコントローラ11002とファイルアクセス制御用CPU11001とファイルアクセス制御用メモリ11004とCPU間通信回路11007とが、ファイバチャネルコントローラに置き換えられた構成である。

【 0 0 3 7 】

(5) ファイルアクセス制御用メモリに格納されているプログラムの一例 (図 5)

図 5 は、CHN110x が有するファイルアクセス制御メモリ 11004 に格納されているプログラムや制御データの一例を示す図である。オペレーティングシステムプログラム 110040 は、プログラム全体の管理や入出力制御に用いられる。LAN コントローラドライバプログラム 110041 は、LAN コントローラ 11002 の制御に用いられる。TCP/IP プログラム 110042 は、LAN 上の通信プロトコルである TCP/IP の制御に用いられる。ファイルシステムプログラム 110043 は、記憶装置に格納されるファイルの管理に用いられる。ネットワークファイルシステムプログラム 110044 は、記憶装置に格納されるファイルを、NAS ホスト 40x に提供するためのプロトコルである NFS や CIFS 等の制御に用いられる。ボリューム制御プログラム 110045 は、ディスクプール 17x 内に設定された記憶領域の一単位である論理ディスクユニット (以下 LU と称する) を複数組み合わせる論理ボリュームを構成するための制御に用いられる。CPU 間通信ドライバプログラム 110046 は、ファイルアクセス制御用 CPU 11001 とディスクアレイ制御用 CPU 11008 間で通信を行うための CPU 間通信回路 11007 の制御に用いられる。

【 0 0 3 8 】

ファイルシステムプログラム 110043 には、

- 1) ファイルを使用する際に、ファイルのオープン処理を実行するファイルオープン処理部 1100431、
- 2) ファイルアクセス要求を受信した際に、アクセス要求に従った処理を実行する要求処理部 1100432、
- 3) ファイルをブロック化し、ブロックのディスク上での格納位置を決定して、ブロックの格納位置を管理するファイル格納管理部 1100433、
- 4) ブロックと、メモリ上に構成されるバッファとの対応づけを管理するバッファ管理部 1100434、
- 5) ファイルを構成するブロックを格納している、ディスク上の記憶領域のアドレスを管理するために用いられる、ファイル格納管理テーブル 1100435、

- 6) オープンされたファイルのファイル名と、そのファイルのファイル格納管理テーブル1100435にアクセスするために用いられるファイルハンドラとを管理するために用いられるファイル名管理テーブル1100436、
- 7) ファイルを構成するブロックの、バッファ内での格納領域を示すバッファアドレスを管理するためのバッファ管理テーブル1100437、
- 8) ファイルの種類、ファイルを生じたアプリケーション、ファイルの生成者の意図などのファイル静的特性と、ファイルのライフサイクルに従い変化するファイルの価値やファイルのアクセス特性などのファイル動的特性とを格納する、ファイル特性情報テーブル1100438、
- 9) ファイルをLU間で移行するための処理を実行する際に用いられるマイグレーション管理部110043A、
- 10) ストレージプール内に構成されたLU毎に、後述するストレージクラスや、LUが存在する記憶装置の識別情報等を登録している、ストレージクラス管理テーブル1100439が含まれる。

【 0 0 3 9 】

(6) ディスクアレイ制御用メモリの構成 (図 6)

図 6 は、ディスクアレイ制御メモリ11009に格納されているプログラムの一例を示す図である。オペレーティングシステムプログラム110090は、プログラム全体の管理や入出力制御に用いられる。ディスクアレイ制御プログラム110091は、ディスクプール17x内にLUを構築したり、ファイルアクセス制御用CPU11001からのアクセス要求を処理するのに用いられる。ディスクプール管理プログラム110092は、ディスクプール17xの構成を、SM13に格納されたディスクプール管理テーブル131の情報を用いて管理するために用いられる。CPU間通信ドライバプログラム110093は、ファイルアクセス制御用CPU11001とディスクアレイ制御用CPU11008との間で通信を行うためのCPU間通信回路11007の制御に用いられる。キャッシュ制御プログラム110094は、CM14に格納されているデータの管理や、キャッシュヒット／ミス判定等の制御に用いられる。DKA通信ドライバプログラム110095は、LUにアクセスする際に、当該LUを構成するディスク170xや171xを制御するDKA12xと通信するために用いられる。

(7) ディスクプール構成 (図 7)

図 7 はディスクプールの構成例を示す図である。

【 0 0 4 0 】

FC ディスクプール 170 には LU0 (51)、LU1 (52) の 2 つの LU が設定されている。LU0 (51) は、DK000、DK010 の 2 台の FC ディスクにより構成されており、DK000、DK010 は RAID1 を構成している。LU1 (52) は DK001、DK002、DK003、DK004、DK005 の 5 台の FC ディスクにより構成されており、これら 5 台の FC ディスクは、4D+1P の RAID5 を構成している。尚、RAID 1 及び RAID5 とは、それぞれ ディスクアレイにおけるデータの配置方法のことであり詳しくは、1988 年に開催された ACM SIGMOD 会議のプロシーディング p.109-116 に掲載されている、D Patterson 外 2 名著、「A Case for Redundant Arrays of Inexpensive Disks (RAID)」に記載されている。RAID1 構成の LU0 では、2 台の FC ディスク DK000 及び DK010 はミラーの関係にある。一方 RAID5 構成の LU1 は、ホスト計算機からアクセスされるファイルのデータが格納されるデータストライプを記憶する一又は複数のディスクと、データストライプに格納されるデータを回復するために用いられるパリティストライプを記憶する一又は複数のディスクから構成される。LU1 は 4 D+1P の RAID5 構成であるが、「4 D+1 P」は 4 つのデータストライプと 1 つのパリティストライプから構成される RAID5 であることを示している。以下同様の表示にて RAID5 構成の LU における、データストライプ数とパリティストライプ数を表示する。

【 0 0 4 1 】

SATA ディスクプール 171 には LU2 (52) が設定されている。LU2 (52) は DK100、DF101、DK102、DK103、DK104、DK110、DK111、DK112、DK113 の 9 台の SATA ディスクにより構成されており、これら 9 台の SATA ディスクは、8D+1P 構成の RAID5 を構成している。

【 0 0 4 2 】

各々のディスクの容量を 140GB とすると、LU0 (51) は 140GB、LU1 (52) は 560GB、LU2 (53) は 1120GB の利用可能な記憶容量を持つ。

【 0 0 4 3 】

尚、各 LU には独立したローカルファイルシステム LFS0 (60)、LFS1 (61)、LFS2 (6

2)が設定、構築されている。

【 0 0 4 4 】

(8) ストレージクラス管理テーブル (図 8)

図 8 は、CHN110xが有するファイルアクセス制御メモリ11004に格納されているストレージクラス管理テーブル1100451の構成例を示している。ストレージクラス管理テーブル1100451は、ファイルアクセス制御用CPUがファイルシステムプログラム110043を実行することによって、SM13上のディスクプール管理テーブル131に格納された情報を参照し、作成したものである。

【 0 0 4 5 】

ディスクプール管理テーブル131については図示しないが、ディスクプール管理テーブル131はSM13に格納されており、ストレージクラス管理テーブル1100451と同様の情報を全CHについて有する。すなわち、あるCHN110xのファイルアクセス制御メモリ11004に格納されているストレージクラス管理テーブル1100451は、ディスクプール管理テーブル131が有する情報のうち、当該CHN110xが使用するLUに関する情報について、ストレージクラスをキーとして配列し直したものである。

以下ストレージクラス管理テーブルの構成を説明する。StorageClass欄(1100451a)には、ストレージクラスを示す情報が格納される。StorageNode#欄(1100451b)には、当該ストレージクラスを構成する記憶装置の識別番号(ストレージノード番号と呼ぶ)が格納される。DiskPOOL#欄(1100451c)には、当該ストレージクラスを構成するディスクプール番号が格納される。LU#欄(1100451d)には、当該ディスクプールに設定されたLUの番号が格納される。LU Type欄(1100451e)には、LUが当該記憶装置の内部(Local)と外部(Remote)のどちらに設定されているかと、LU内にファイルシステムが設定されているかいないかを示す情報が格納される。即ち、LUが記憶装置内に存在する場合には“Local”が、LUが他の記憶装置内に存在する場合には“Remote”がLU Type欄に登録され、LUにファイルシステムが構築されている場合には“File”が、LUにファイルシステムが構築されていない場合には“Block”がLU Type欄に登録される。RAIDConf.欄(1100451f)には、当該LUを構成するディスクアレイのRAIDレベルと、パリティグループ内のデー

タレコードとパリティレコードの数等、ディスクアレイの構造を示す情報が格納される。Usable Capacity欄(110045lg)とUsedCapacity欄(110045lh)には各々、当該LUの総記憶容量と使用中の記憶容量を示す情報が格納される。

【 0 0 4 6 】

ストレージクラスとは、データの格納用途別に設けられた、記憶領域の階層的な属性であり、本実施形態では、OnLine Storage、NearLine Storage、Archive Storageの3つの属性を定義している。また、OnLine Storageについては、副属性としてPremium、Normalを定義している。OnLine Storageは、オンラインアクセス中のファイルや、生成途上のファイル等、頻繁にアクセスされるファイルのデータを格納するのに適するLUに設定される属性である。特に、Premiumは早いレスポンスを要求されるデータを格納するのに適したLUに設定される属性である。NearLine Storageは、あまり頻繁には使用されないが、まれにアクセスされるファイルのデータを格納するのに適するLUに設定される属性である。Archive Storageは、ほぼアクセスされることはなく、長期保存のためのファイルのデータを格納するのに適するLUに設定される属性である。

【 0 0 4 7 】

図8では、記憶装置1 (STR0と呼ぶ) が有するFCディスクプール170にOnLine Storage(Premium)クラスのLU0(50)とOnLine Storage(Normal)クラスのLU1(51)が存在していることを示している。また、同じ記憶装置1(STR0)内の、SATAディスクプール171には、NearLine StorageクラスのLU2(52)が存在する。また、別の記憶装置(STR1)には、SATAディスクプールにArchive StorageクラスのLU3(53)が存在する。なお、異なる記憶装置にディスクプールを構築する例は後述する。

(9) ファイル名管理テーブル (図9)

図9には、ファイルアクセス制御用メモリ11004に格納されている、ファイル名管理テーブル1100436の一例を示す。ファイル名管理テーブルはファイルシステム毎に用意されるテーブルであり、ファイル名とファイルハンドラが、容易に検索できるように、木構造で格納されている。ファイルがNASホスト40xからアクセスされる際には、NASホストからCHNが受信するアクセス要求にファイル名が含まれている。CHNはこのファイル名を用いてファイル名管理テーブル11004を検索

し、ファイル名に対応するファイルハンドラを取得して、このファイルハンドラと対応付けられているファイル格納管理テーブル1100435を参照することができる。

【 0 0 4 8 】

なお、ファイル名管理テーブルは、ファイル名管理テーブルに対応するファイルシステムが構築されているLU上に格納されており、必要な時にメモリに読み込まれてファイルアクセス制御用CPUによって使用される。

(10) ファイル格納管理テーブル (図 1 0)

図 1 0 はファイル格納管理テーブル1100435及びバッファ管理テーブル1100437の一例を示す図である。ファイル格納管理テーブルは、ファイル毎にファイルアクセス制御用メモリに備えられており、ファイルの格納アドレスを管理するテーブルである。ファイル格納管理テーブルは、ファイルを示すファイルハンドラを指定することによって参照される。

【 0 0 4 9 】

ファイル特性情報管理テーブル欄には、対応するファイルについてのファイル特性情報管理テーブル1100438を参照するためのポインタが格納されている。大きさはファイルのバイト単位の大きさを示す。ブロック数は、ファイルを論理ブロックと呼ぶブロックに分割して管理する際に使用している論理ブロックの数を表す。また、ファイルが格納される論理ブロックごとに、その論理ブロックについてのバッファ管理テーブル1100437へのポインタが格納されている。

【 0 0 5 0 】

バッファ管理テーブル1100437は、各論理ブロック毎に1つ存在し、以下の内容が格納されている。ハッシュリンク欄にはバッファが有効かどうかすばやく判定するためのハッシュ表へのリンクポインタが格納されている。待ち行列リンク欄には待ち行列を形成するためのリンクポインタが格納されている。フラグ欄にはバッファの状態すなわち、有効なデータが格納されているかどうか、バッファは使用されているかどうか、バッファの内容がディスクに未反映となっているかどうか、等を示すフラグが格納されている。機器番号欄には論理ブロックが格納される記憶装置の識別子及びLUの識別子が格納されている。ブロック番号欄には

機器番号で示された記憶装置内での、論理ブロックの格納位置を示すディスクアドレス番号が格納されている。バイト数欄には論理ブロックに格納された有効なデータのバイト数が格納されている。バッファサイズ欄には本バッファのバイト単位の大きさが格納されている。バッファポインタ欄には物理バッファメモリへのポインタが格納されている。

【 0 0 5 1 】

なお、ファイル格納管理テーブルは、対応するファイルが格納されるLU内に格納され、必要な時にメモリに読み込まれて使用される。

(11) ファイル特性情報管理テーブル (図 1 1)

図 1 1 は、ファイルアクセス制御用メモリ 11004 に格納されているファイル特性情報管理テーブル 1100438 の一例を示す図である。ファイル特性情報管理テーブルは、静的特性情報と動的特性情報を格納する。静的特性情報は、ファイルを構成した際に決定され、その後も引き継がれる情報である。もちろん静的特性情報を意図的に変更することもできるが、その時以外に変更されることはない。動的特性情報は、ファイルが作成された後、時間の経過に従って逐次変更されていく情報である。

(12) 静的特性情報

静的特性情報には、ファイル情報カテゴリとポリシーカテゴリがある。

【 0 0 5 2 】

ファイル情報カテゴリには、ファイルの基礎情報が含まれる。ファイル情報カテゴリにおいて、ファイル種類は、テキストファイル、ドキュメントファイル、画像ファイル、動画ファイル、音声ファイル等、当該ファイルの種類を示す。アプリケーションは、当該ファイルを生成したアプリケーションを示す。作成日は、最初にファイルが生成された日付を示す。作成日としてファイルが作成された日付だけでなく作成時間も登録しても良い。所有者はファイルを作成したユーザ名を示す。アクセス識別子は当該ファイルのアクセス許可範囲を示す。

【 0 0 5 3 】

ポリシーカテゴリは、ファイルを生成したユーザ又はアプリケーションによって設定される情報であり、ファイルの格納条件等について、ユーザ又はアプリケ

ーションにより指定される情報である。初期ストレージクラスは、ファイルを最初に記憶装置に格納する際に、そのファイルを格納すべきLUのストレージクラスを示す情報である。資産価値タイプは、当該ファイルの資産価値を示す。ライフサイクルモデルは、あらかじめ定義されたライフサイクルモデルのうち、当該ファイルに適用されるモデルを示す。マイグレーションプランは、あらかじめ定義されたファイルの移行（以下、マイグレーションと称する。）に関するプランのうち、当該ファイルに適用されるプランを示す。

【0054】

ここで、資産価値とは、ファイルがどのような重要性や価値を持っているかを指定する属性である。資産価値としては例えば、「特別重要」、「重要」、「普通」等の属性を指定できる。例えば、「重要」以上の属性のファイルはPremium属性のストレージクラスに属するLUに格納する等、資産価値をストレージクラス選定の際の補足基準として用いたり、ライフサイクルモデルが指定されていないときのストレージクラス選定の基準として使用することができる。

【0055】

以下本実施形態の説明には、「重要」以上のファイルは「Premium」クラスのストレージクラスに属するLUに格納することと仮定する。勿論、本発明はこの様な仮定に限定されることはなく、ファイルを格納するLUのストレージクラスを選択する際に、異なる基準を用いることも可能である。

【0056】

ライフサイクルとは、データが生成される時期を誕生、データが更新され又利用される時期を成長、データが更新されることは少なく、参照されることが主となる時期を成熟、そしてデータがやがて使用されなくなり保管されるようになる時期を老齢といった様に、ファイルの使用状況が時間経過によって移り変わってゆくのを、人の生涯に見立て、名付けられたものである。ライフサイクルモデルとは、ファイルがどのようなライフサイクルを送るかを定義したものである。最も汎用的な定義の方法としては、ファイル生成後の経過時間で定義する方法がある。一例としては、更新が頻発する「成長期」もしくは「更新期」を1ヶ月、参照が主体となる「成熟期」もしくは「参照期」を1年、それ以降を「老齢期」も

しくは「保管期」と定義する方法がある。以下、この定義を「モデル1」と称し、以下の説明に用いる。ライフサイクルモデルの時間間隔を変更したり、さらなる分解能をもった期間を定義することで、色々なライフサイクルモデルを定義し、複数のライフサイクルモデルからあるライフサイクルモデルを選択して使用することができる。特定のファイル種類に特定のライフサイクルモデルを適用したり、アプリケーション毎にそのアプリケーションが生成したファイルに対して適用するライフサイクルモデルを用意してもよい。なお、ライフサイクルの名称には、人の生涯に対応した「成長期」「成熟期」「老齢期」という表現を用いる場合や、ファイル挙動に基づく「更新期」「参照期」「保管期」という表現を用いる場合がある。以下本実施形態では挙動を明確にするため後者を用いることにする。

【0057】

マイグレーションプランとは、ライフサイクルに従い、ファイルをどのようなストレージクラスのLUに移動していくかを定義したものである。一例として、「更新期」のファイルをOnLine Storageクラスの、「参照期」のファイルをNearLine Storageクラスの、「保管期」のファイルをArchive StorageクラスのLUに格納する方法がある。以下、この定義を「プラン1」と称し、以下の説明で用いる。このプラン以外にも、「更新期」のファイルはOnLine Storage (premium)クラスに、「参照期」のファイルはOnLine Storage(Normal)クラスに格納するよう定義したり、「保管期」のファイルもNearLine Storageクラスに格納したままにする等、いろいろなプランを定義し、その中からプランを選択して使用できるようにすることができる。特定のファイル種類のファイルに特定のマイグレーションプランを適用したり、アプリケーション毎にそのアプリケーションが生成したファイルに適用するマイグレーションプランを用意してもよい。

(13)動的特性情報

動的特性情報には、アクセス情報カテゴリと、ライフサイクル情報カテゴリがある。

【0058】

アクセス情報カテゴリには、ファイルのアクセス統計情報が含まれる。アクセ

ス情報カテゴリにおいて、タイムスタンプはファイルが最後に読み書きされた日付や時刻や、ファイルのファイル格納管理テーブルが最後に更新された日付や時刻を示す。アクセスカウントはファイルに対する総アクセス回数を示す。リードカウント、ライトカウントは各々、ファイルへのリード回数、ライト回数を示す。リードサイズ、ライトサイズはファイルのリード時とライト時それぞれにおける、データ転送サイズの平均値を示す。リードシーケンシャルカウント、ライトシーケンシャルカウントは、リード時とライト時それぞれにおいて、連続する複数の2回のアクセス間にアドレス連続性、すなわちシーケンシャル性がある場合の回数を示す。

【 0 0 5 9 】

ライフサイクル情報カテゴリには、ファイルのライフサイクルに関連する情報が含まれる。ライフサイクル情報カテゴリにおいて、現行ライフサイクルは、更新期、参照期、保管期等、現時点でのファイルのライフサイクルにおける位置づけを示す。現行ストレージクラスは、現時点においてファイルが格納されているLUが設定されているストレージプールのストレージクラスを示す。

【 0 0 6 0 】

以上、図11には、ファイル特性情報の一例を示したが、これら以外にも色々な特性情報を定義してファイル特性情報管理テーブル1100438に格納することができる。また、これらの特性情報の一部のみを必要に応じて用いる実施形態とすることもできる。

(14) ファイル初期配置：ファイルオープン処理

次に、ファイルを記憶装置に最初に格納する初期配置処理の際に行われる、ファイルオープン処理について説明する。

【 0 0 6 1 】

NASホスト0(400)がファイルabc.docを生成したとする。

【 0 0 6 2 】

NASホスト0(400)は、CHN0(1100)に対してファイルabc.docのオープン要求を発行する。オープン要求には、ファイルを識別するための識別情報としてファイル名が含まれる。今回のオープン処理は、ファイルの新規格納のために実行される

ので、NASホスト0(400)は、ファイル特性情報の静的特性情報として、ファイル情報カテゴリとポリシーカテゴリに含まれる以下の情報をオープン要求と共にCHNO(1100)へ送信する。送信される情報は、ファイル情報カテゴリに含まれる情報として、ファイル種類「ドキュメント」、ファイルを生成したアプリケーション「X Y Zワード」、アクセス識別子「-rw-rw-rw-」と、ポリシーカテゴリに含まれる情報として、初期ストレージクラス「指定なし」、資産価値タイプ「重要」、ライフサイクルモデル「モデル1」、マイグレーションプラン「プラン1」である。

【 0 0 6 3 】

CHNO(1100)は、LANコントローラ11002経由でNASホストからファイルのオープン要求を受信し、ファイルアクセス制御用CPU11001はファイルシステムプログラム110043を実行する。

【 0 0 6 4 】

ファイルシステムプログラム110043が実行されると、ファイルアクセス制御用CPU11001の制御により、受信されたオープン要求は、ファイル名のディレクトリ情報に基づき、ローカルファイルシステムL F S 0 (60)へのアクセス要求であることが特定される。そして、ファイルオープン処理部1100431は、L F S 0 (60)のファイル名管理テーブル1100436を参照し、abc.docを検索する。その結果、abc.docがまだファイル名管理テーブル1100436には存在しない新たに格納されるファイルであることが判明するので、ファイルオープン処理部1100431は、ファイル名管理テーブルにabc.docを登録しファイルハンドラを割り当てる。

【 0 0 6 5 】

次にファイル格納管理部1100433がファイルabc.docに割り当てられたファイルハンドラと対応させてファイル格納管理テーブル1100435を作成する。

【 0 0 6 6 】

次にファイル格納管理部1100433が、ファイル特性情報管理テーブル1100438を生成して、ファイル格納管理テーブル1100435に関連づけた上で（即ちファイル格納管理テーブルにファイル特性情報管理テーブルへのポインタを格納した上で）、NASホスト0から取得したファイル特性情報の静的特性情報と、ファイルの作

成日、所有者をファイル特性情報管理テーブル1100438に格納する。そして、このファイルが属するファイルシステムが構築されているLUに、ファイル格納管理テーブル及びファイル特性情報管理テーブルが書き込まれる。

【 0 0 6 7 】

その後、CHN0(1100)はファイルハンドラをNASホスト0に返しオープン処理が終了する。

(15)ファイル初期配置：データライト処理

次にファイルの初期配置処理の際に実行されるデータライト処理について説明する。

【 0 0 6 8 】

NASホスト0(400)は、ファイルabc.docのデータを記憶装置1に格納すべく、オープン処理で獲得したファイルハンドラを用いてライト要求をCHN0(1100)に発行する。

【 0 0 6 9 】

CHN0 (1100) では、ライト要求を受信するとファイルアクセス制御用CPU11001がファイルシステムプログラム110043を実行し、オープン処理の際と同様の方法によってこのライト要求がローカルファイルシステムL F S 0 (60)へのアクセス要求であることを特定する。

【 0 0 7 0 】

ファイルシステムプログラム110043の要求処理部1100432は、受信したアクセス要求内に含まれる情報に基づいてこのアクセス要求がライト要求であることを解釈し、ライト要求内で指定されているファイルハンドラを用いて、このファイルハンドラに対応付けられているファイルのファイル格納管理テーブル1100435を獲得する。

【 0 0 7 1 】

次にファイル格納管理部1100433は、データを格納するのに必要なバッファを確保するとともに、ファイルのディスク上での格納位置を決定する。

【 0 0 7 2 】

格納位置の決定に際しては、ファイル格納管理部1100433は、ファイル特性情

報管理テーブル1100438の静的特性情報を参照する。ここで、ライト要求の対象となっているファイルabc.docのライフサイクルモデルは「モデル1」であり、受信したライト要求はファイルの初期配置の際のアクセス要求であるからファイル生成後1ヶ月以内のアクセスなので、ファイル格納管理部1100433はファイルabc.docの現行ライフサイクルは「成長期」とであると特定する。また、初期ストレージクラスは「指定なし」であり、資産価値タイプは「重要」であるので、ファイル格納管理部は、ファイルabc.docを格納すべきストレージプールのストレージクラスとして「OnLine Storage(Premium)」を選択する。

【 0 0 7 3 】

次にファイル格納管理部1100433は、ストレージクラス管理テーブル1100439を参照し、ストレージクラス「OnLine Storage(Premium)」に該当する、ストレージノード「STR0（すなわち自記憶装置1）」、DiskPool#「FCディスクプール1700」、LU#「LU0（すなわちローカルファイルシステムはL F S 0）」で特定されるLUにファイルを格納することを決定する。そしてファイル格納管理部1100433は、適切なアルゴリズムの下、ファイル内のデータを一又は複数の論理ブロックに分割し、論理ブロック毎にLU0上の格納アドレスを決定し、バッファ管理テーブル1100437を生成して決定された格納アドレスを登録し、ファイル格納管理テーブル1100435のバッファ管理テーブルエントリに、生成されたバッファ管理テーブルへのポインタを格納する。更に、ファイル格納管理部1100433は、ファイル格納管理テーブル1100435の残りのエントリに情報を格納する。なお、本実施形態においては、ファイル格納管理テーブルのリンク先のエントリは全てNULLとする。

【 0 0 7 4 】

そして、ファイル格納管理部1100433は、ファイル特性情報管理テーブル1100438の動的特性情報のライフサイクル情報カテゴリについて、現行ライフサイクルを「更新期」と、現行ストレージクラスを「OnLine Storage(Premium)」と設定する。また、ファイル格納管理部1100433は、動的特性情報のアクセス情報カテゴリに含まれる情報について、適切に計算した上で、ファイル特性情報管理テーブルに登録する。

【 0 0 7 5 】

要求処理部1100432は、受信したライト要求に従った処理を実行し、LANコントローラドライバプログラム110041、TCP/IPプログラム110042、ネットワークファイルシステムプログラム110044がファイルアクセス制御用CPU11001によって実行された結果、NASホスト0(400)からライトデータがCHN0(1100)に転送され、ファイルアクセス制御用メモリ11004のバッファに一旦格納される。その後、CPU間通信ドライバプログラム110046がファイルアクセス制御用CPU11001によって実行され、この結果適切なタイミングでディスクアレイ制御用CPUにライト要求が転送される。ライト要求を受信したディスクアレイ制御用CPU11008は、CM14にライトデータを一旦キャッシュし、NASホスト0(400)からのライト要求に対して完了応答を返信する。

【 0 0 7 6 】

その後、適切なタイミングで、LU0を構成するディスクを制御するDKA120の制御の下、当該ディスクにライトデータが格納される。

【 0 0 7 7 】

以上のように、ファイルの静的特性情報に基づき、適切なストレージクラスに属する記憶領域にファイルを初期配置することができる。

(16) ファイルマイグレーション処理 (図 1 2)

次にファイルのマイグレーション処理について説明する。

【 0 0 7 8 】

ファイルシステムプログラム110043のマイグレーション管理部110043Aは、あらかじめ設定されたタイミングで、ファイルアクセス制御用CPUによって起動される。

【 0 0 7 9 】

そして、マイグレーション管理部110043Aは、あらかじめマイグレーション処理の対象として設定されているローカルファイルシステムについて、そのファイルシステムに含まれるファイルのファイル特性管理テーブルを参照することにより、マイグレーションの対象となるファイルが存在するかどうか調査する。以下、ファイルabc.docがマイグレーション処理の対象となる場合について、具体的

に説明する。

【 0 0 8 0 】

マイグレーション管理部110043Aは、abc.docのファイル特性情報管理テーブル1100438を参照し、作成日と現在の日付及び時刻を比較する。既に作成日から1ヶ月が経過していたとすると、静的特性情報のライフサイクルモデルが「モデル1」であり、「更新期」の期間である1ヶ月間が経過していることから、既に現行ライフサイクルが「更新期」から「参照期」に変化しているということをマイグレーション管理部110043Aが認識する。

【 0 0 8 1 】

更にマイグレーションプランが「プラン1」であることから、ファイルを、ストレージクラスが「OnLine Storage (Premium)」であるLUから「NearLine Storage」であるLUに移行しなくてはならないことを、マイグレーション管理部110043Aは認識する。

【 0 0 8 2 】

そこでマイグレーション管理部110043Aは、ストレージクラス管理テーブル1100439を参照し、「NearLine Storage」クラスに属する、ストレージノード「STR0 (すなわち自記憶装置1)」、DiskPool#「SATAディスクプール1710」、LU#「LU2 (すなわちローカルファイルシステムはL F S 2)」で指定されるLUに、当該ファイルを移動することを決定する。

【 0 0 8 3 】

次にマイグレーション管理部110043Aは、ファイル特性情報管理テーブル1100438の動的特性情報の、現行ライフサイクルを「参照期」に、現行ストレージクラスを「NearLine」にそれぞれ変更する。

【 0 0 8 4 】

そして、マイグレーション管理部110043Aは、ファイルabc.docを記憶装置STR0 (1)内部で管理する際に使用するための、一意なファイル名（ここではFILE00001とする）を定義する。

【 0 0 8 5 】

ファイルオープン処理部1100431は、L F S 2 (60)のファイル名管理テーブル1

100436を参照し、FILE00001のファイル名が未だファイル名管理テーブル1100436に登録されていないことを確認し、登録されていなければ、ファイル名管理テーブル1100436にファイル名FILE00001に登録し、ファイルハンドラを割り当てる。

【 0 0 8 6 】

次にファイル格納管理部1100433が、ファイルFILE00001に割り当てられたファイルハンドラと対応させて、ファイル格納管理テーブル1100435とファイル特性情報管理テーブル1100438を生成する。生成されたファイル特性情報管理テーブル1100438には、abc.docのファイル特性情報管理テーブルに登録されている内容と同じ内容が格納される。そして、ファイル格納管理部1100433は、FILE00001が格納されるLUに、FILE00001のファイル格納管理テーブルとファイル特性情報管理テーブルを書き込む。

【 0 0 8 7 】

次にファイル格納管理部1100433は、FILE00001のデータを格納するのに必要なバッファ領域を確保するとともに、このファイルを格納するためのLU2内の記憶領域（若しくは格納位置）を決定する。そして前述のデータライト処理の際と同様の方法で、ファイル格納管理部1100433は、バッファ管理テーブル1100438を生成して決定された格納位置に登録し、ファイル格納管理テーブル1100435のバッファ管理テーブルエントリに生成されたバッファ管理テーブルへのポインタを登録する。尚、L F S 2 に格納されるファイルFILE00001のファイル格納管理テーブル1100435のリンク先エントリは、全てNULLであるものとする。

【 0 0 8 8 】

また、ファイル格納管理部1100433は、図12に示すように、L F S 0 のabc.docのファイル格納管理テーブル1100435のリンク先ノード名をSTR0、リンク先FS名をLFS2、リンク先ファイル名をFILE00001に変更する。

【 0 0 8 9 】

次に要求処理部1100432は、abc.docのデータをLU0を構成しているディスクからファイルアクセス制御用メモリ11004内のバッファに読み込む。そして、ファイルアクセス制御用メモリ11004内のバッファに読み込まれたこのデータを、LU2を構成しているディスクに書き込まれるFILE00001のデータとして、上述のファイ

ル格納管理部1100433によって決定され、バッファ管理テーブルに登録されているバッファ内の格納領域に、要求処理部1100432が書き出す。

【0090】

そして、ファイル格納管理部1100433は、L F S 0 のabc.docのファイル格納管理テーブル1100435に登録されているポインタから参照可能なバッファ管理テーブルを全て解放し、これらのバッファ管理テーブル内のエントリにNULLに登録する。

【0091】

バッファに格納されたFILE00001のデータは、適切なタイミングで前述の初期配置処理におけるデータライト処理と同様の手順により、記憶装置1のCM14を経由て、LU2に格納される。これでマイグレーション処理が終了する。

【0092】

以上のように、本実施形態によれば、ファイルのマイグレーションプランに基づき、ファイルのライフサイクルを考慮した適切なストレージクラスの記憶領域に、ファイルをマイグレーションすることができる。

【0093】

本実施形態によれば、ホスト計算機やホスト計算機で実行されるアプリケーションに依存することなく、ストレージクラスという概念に基づいてファイルを格納するためのLUを選択し、またファイルを格納するLUを変更することができる。この結果、ホスト計算機に依存することなく、コスト効果の高い記憶階層、すなわち、特性の異なる複数の記憶領域、を有する記憶装置を実現することができる。

【0094】

また、ファイル単位でデータの移行が実行されるので、ファイルI/Oインタフェースを用いれば、ファイルの移行後も、複数のホスト計算機から同じファイルにアクセスすることができる。

【0095】

更に、ファイルの種類やファイルを生成したアプリケーションの種類、ファイル生成者の意図（ポリシー）といったファイルの静的特性や、ファイルのライフ

サイクルや、価値、アクセス特性の変化といったファイルの動的特性に基づいたファイルベースの階層記憶制御を実行することができる。

【 0 0 9 6 】

[実施形態 2]

(1) システム構成の一例 (図 1 3)

次に第 2 の実施形態のシステム構成の一例を図 1 3 を用いて説明する。本実施形態においては、第 1 実施形態の記憶装置 1 (以下 STR0 と称する) と、別の記憶装置 1a (以下 STR1 と称する) をネットワークで接続したシステムにおいて、記憶装置間で階層記憶制御が実行される。

【 0 0 9 7 】

図 1 3 において、記憶装置 STR1 (1a) は、記憶装置 STR0 (1) に LAN20 を介して接続される他の記憶装置であり、その他のシステム構成要素は図 1 と共通である。

【 0 0 9 8 】

STR1 (1a) において、NCTL0 (1100a) 及び NCTL1 (1101a) は NAS コントローラ、ディスクプール 0 (170a) は NCTL0、NCTL1 に接続されたディスクプールである。

【 0 0 9 9 】

NAS コントローラ NCTLx は、図 4 に示す第 1 実施形態の CHN1100 の構成のうち、SM I/F 制御回路 11005 及び CM I/F 制御回路 11006 に代わり、ディスクプール 0 1700a を接続するための FC コントローラ 11010a を備える。また、NAS コントローラ NCTLx は、キャッシュメモリ CM14a を NAS コントローラ内に有し、またキャッシュメモリ CM14a のための制御回路であるデータ転送回路 11011a を備える。更に NAS コントローラ NCTLx は、データ転送回路 11011a を有し、データ転送回路を介して NAS コントローラ 1100a、NAS コントローラ 1101b は相互に接続されている。図 13 では NAS コントローラ NCTL1 (1101a) の構成の詳細は図示していないが、NAS コントローラ 1101a は、NAS コントローラ 1100a と同様の構成を有する。尚、第一の実施形態において説明した CHN1100 と同じ番号が振られた構成要素は CHN1100 の該当要素と同様の構成を有し、同様の機能を有する。

【 0 1 0 0 】

STR1 は、STR0 に比べて小規模かつ低価格な記憶装置であるとする。また、図 13

に示す通り、STR0のCHN0とSTR1のNCTL0はLAN20により接続されている。

(2)他の記憶装置へのファイルの移行処理

次に本実施形態の動作を説明する。

【0101】

記憶装置1(STR0)のCHN0(1100)は、LAN20に種類の異なる記憶装置1a(STR1)が接続していることを認識する。異なる記憶装置の認識は、あらかじめ管理者から指定された情報に基づき行う方法や、LAN20のネットワークセグメントに認識のためのコマンドを同報送信することにより反応した装置の有無により行う方法等がある。STR0のCHN0はSTR1の構成を把握するため、自身がイニシエータとなり、情報を採取するためのコマンドをSTR1に発行する。このコマンドに対するSTR1からの応答には、STR1が有するディスクプールの種類やLUの構成が含まれているので、この応答を参照することによって、CHN0は、STR1がSATAディスクプール170aを所持し、このディスクプール170a内に、15D+1PのRAID5構成を有し2100GB容量の大容量で低コストなFile TypeのLUを所持していることを認識することができる。そして、STR0のCHN0は、STR1が有するLUを、リモートLU、即ち他の記憶装置STR1(1a)に存在するが自記憶装置STR0(1)が管理する一つのLUとして、管理することを決定する。

【0102】

CHN0はSTR1が有するLUにLU3の番号を割り当て、またこのLU内に構築されるファイルシステムにはリモートファイルシステムRFS3の番号を割り当てる。また、このLUは大容量低コストなディスクプールに存在するので、このLUのストレージクラスを「Archive Storage」に設定する。そしてCHN0のディスクアレイ制御用CPU11008の制御によって、上述のディスクプールの種類、LUの構成、LU番号、ストレージクラス等、STR1に存在するLU3に関する情報が、記憶装置1(STR0)が有するSM13のディスクプール管理テーブル131に格納される。記憶装置1が有するCHNは、ファイルシステムプログラム110043をファイルアクセス制御用CPU11001が実行することによって、ディスクプール管理テーブル131を参照し、ディスクプール管理テーブル131からLU03に関する情報をコピーすることによって、LU03についての情報をファイルアクセス制御用メモリ内のストレージクラス管理

テーブル1100451に登録することができる。

【 0 1 0 3 】

以下、第1実施形態で説明したように、NASホスト0(400)がファイルabc.docをCHN0経由でSTR0のLU0に格納し、その後CHN0の制御によってファイルabc.docがSTR0のLU2に移行されたと仮定して、本実施形態において、ファイルabc.docを更に他の記憶装置STR1内のLU3に移行するために実行される処理について、第1実施形態との相違点のみを説明する。

【 0 1 0 4 】

第1実施形態で説明した通り、ファイルabc.docは、図11、図12に示すように、現行ライフサイクルが「参照期」で、現行ストレージクラスは「NearLine Storage」で、STR0のSATAディスクプールに構築されたLU2上のLFS2にFILE00001という名称でデータ部が格納されている。ファイル名「abc.doc」が登録されているファイル名管理テーブル1100436と、ファイルabc.docに関するファイル格納管理テーブル1100435は、LFS0に存在する。即ち、LFS0用のファイル名管理テーブル110436及びLFS0内のファイル用のファイル格納管理テーブルに、abc.docに関する情報が登録されている。一方、ファイル特性管理情報管理テーブル1100438はLFS0とLFS2の両方に存在している。尚、ファイルabc.docのデータ部はLFS2が構築されているLU2にマイグレーション済みなので、LFS0が構築されているLU0にはabc.docのデータ部は存在しない。

【 0 1 0 5 】

STR0のマイグレーション管理部110043Aは、abc.docのファイル特性情報管理テーブル1100438を参照し、作成日と現日付及び時刻を比較する。既にマイグレーション後1年が経過していたとすると、abc.docの静的特性情報のライフサイクルモデルは「モデル1」であるから、「参照期」の期間である1年間が経過しており、既に現行ライフサイクルが「参照期」から「保管期」に変化しているということを、マイグレーション管理部110043Aは認識する。そして、abc.docのマイグレーションプランが「プラン1」であることから、マイグレーション管理部110043Aは、このファイルをストレージクラスが「NearLine Storage」であるLUからストレージクラスが「Archive Storage」であるLUに移行しなくてはならない

ことを認識する。

【0106】

次にマイグレーション管理部110043Aは、ストレージクラス管理テーブル1100439を参照し、「Archive Storage」クラスに属するLU3を選択して、LU3にファイルabc.docを移動することを決定する。LU3は、ストレージノードが「STR1（すなわち他の記憶装置1a）」、DiskPool#が「SATAディスクプール」、LU Typeが「Remote File」という属性を備える。

【0107】

次にマイグレーション管理部110043Aは、abc.docのファイル特性情報管理テーブル1100438の動的特性情報の、現行ライフサイクルを「保管期」に、現行ストレージクラスを「Archive」にそれぞれ変更する。

【0108】

次にマイグレーション管理部110043Aは、記憶装置STR0(1)内でファイルabc.docを管理するための一意なファイル名（ここではSTR1-FILE00001とする）を定義する。

【0109】

マイグレーション管理部110043Aは、自身がNASホストのごとく動作し、STR1に対してファイルSTR1-FILE00001のオープン要求を発行する。今回のオープン処理は、STR1にとってはファイルの新規格納のために実行されるオープン処理である。そこでSTR0は、ファイルabc.docの静的特性情報として、STR0がファイル特性情報管理テーブル内に所持している情報を、オープン要求内に含めて、STR1に送信する。ただし、静的特性情報のうち初期ストレージクラスのみ「Archive」に変更して送信することで、STR0はSTR1に対して、ファイルSTR1-FILE00001を最初からArchiveストレージクラスに格納するよう明示的に指定する。

【0110】

STR1のNCTL0は、LANコントローラ11002a経由でオープン要求を受信し、ファイルアクセス制御用CPU11001aはファイルシステムプログラム110043aを実行する。

【0111】

ファイルシステムプログラム110043aが実行されると、第1実施形態と同様に

、受信したオープン要求がローカルファイルシステム R F S 3 へのアクセス要求であることが特定され、ファイルアクセス制御用CPU11001aの制御によって、ファイルアクセス制御用メモリ11004a内のファイル名管理テーブル1100436aにSTR1-FILE00001が登録されてファイルハンドラが割り当てられ、ファイルアクセス制御用メモリ11004a内にファイル格納管理テーブル1100435a及びファイル特性情報管理テーブル1100438aが作成され各テーブル内に登録されるべき情報が設定される。そして、NCTL0はSTR1-FILE00001に対して割り当てられたファイルハンドラを、CHN0のマイグレーション管理部110043Aに宛てて送信し、オープン処理が終了する。

【 0 1 1 2 】

次にSTR0のマイグレーション管理部は、第1実施形態のデータライト処理におけるNASホストと同様、オープン処理でSTR1のNCTL0から取得したファイルハンドラを有するライト要求をSTR1に対して発行し、abc.docの実データ（即ちFILE00001の実データでもあるデータ）をファイルSTR1-FILE00001の実データとしてライトするよう要求する。

【 0 1 1 3 】

STR1のファイル格納管理部1100433aは、ライトデータを保存するために必要なバッファ領域を確保し、ファイルの実データのディスク上での格納位置を決定し、又、STR0から受信したライトデータをバッファに格納する。

【 0 1 1 4 】

格納位置の決定に際しては、ファイル格納管理部1100433aによってファイル特性情報管理テーブル1100438aの静的特性情報が参照される。ここで、ファイルSTR1-FILE00001のライフサイクルモデルは「モデル1」であり、ファイル生成後1年1ヶ月以上経過しているので、STR1-FILE00001の現行ライフサイクルは「保管期」であるとファイル格納管理部1100433aは特定する。また、初期ストレージクラスはSTR0から指定された「Archive Storage」であるとファイル格納管理部1100434aは特定する。

【 0 1 1 5 】

そしてファイル格納管理部1100433aは、ファイル特性情報管理テーブル110043

8aの動的特性情報のライフサイクル情報カテゴリについて、現行ライフサイクルを「保管期」と、現行ストレージクラスを「Archive Storage」と設定する。更にファイル格納管理部1100433aは、ファイルSTR1-FILE00001に関するアクセス情報を適切に計算し、ファイル特性情報管理テーブルのアクセス情報カテゴリ内の情報更新する。尚、STR1-FILE00001のファイル格納管理テーブル1100435aのリンク先エントリには全てNULLが登録されるものとする。

【 0 1 1 6 】

その後、NCTL0の制御により、適切なタイミングで、LU3を構成するディスクにファイルSTR1-FILE00001のデータ部分が格納される。

【 0 1 1 7 】

以上でSTR1におけるライト処理が終了し、処理がSTR0に戻る。

【 0 1 1 8 】

STR0のファイル格納管理部1100433は、L F S 2 のFILE00001のファイル格納管理テーブル1100435のリンク先ノード名をSTR1、リンク先FS名をLFS3、リンク先ファイル名をSTR1-FILE00001に変更する。そして、ファイル格納管理部1100433は、FILE00001のファイル格納管理テーブル1100435に登録されているポインタから参照されるバッファ管理テーブルを全て開放し、これらのバッファ管理テーブルのエントリをNULLとする。

【 0 1 1 9 】

以上により、ファイルabc.docのデータ部の実態はSTR0のL F S 2 にあったFILE00001から、STR1のR F S 3 にあるSTR1-FILE00001へ移動したことになる。

【 0 1 2 0 】

以降、NASホストからファイルabc.docに対するアクセス要求が発行された場合には、STR0のCHNが一旦L F S 0 のabc.docのファイル格納管理テーブルを参照してリンク先ノード名、FS名、ファイル名を取得し、取得したリンク先の識別情報（即ち、STR0、LFS2、FILE00001）に基づいて、L F S 2 のFILE00001のファイル格納管理テーブルを参照する。そしてSTR0のCHNは更に、LFS 2 のFILE00001のファイル格納管理テーブルから再びリンク先ノード名、FS名、ファイル名を取得し、取得したリンク先の識別情報（即ち、STR1、LFS3、STR1-FILE00001）を指定し

たアクセス要求を、STR1のNCTLへ発行することによって、STR1のR F S 3のSTR1-FILE00001にたどり着き、STR1のNCTLを介してabc.docのデータ部にアクセスできる。

【 0 1 2 1 】

このように本実施形態においては、STR1が有するLU3内に移行されたファイルにアクセスするためには、複数のファイル格納管理テーブルを参照しなければならないため、アクセス速度は若干低下する。しかしSTR1が有するLU3に格納されるファイルは、現行ライフサイクルが「保管期」のファイルであり、ほとんどアクセス要求がこないファイルであるので、実用上は問題ない。仮に、「保管期」のファイルのデータに対してホスト計算機からアクセス要求が発行された場合であっても、Archiveクラスに属するファイルとはいえ磁気ディスク内に格納されているため、従来の様にテープにファイルが格納されている場合と比べれば、テープ制御のための多大なアクセス時間を必要とせず、一旦テープからディスクにデータを移動しなおす必要もなく、データが記憶されているディスク上の記憶位置からそのままリアルタイムでデータを取り出すことができる。

【 0 1 2 2 】

以上のように、本実施形態によれば、ライフサイクルに従いファイルの格納位置を決定しているので、ライフサイクルが老いた「保管期」のファイルには保管に適したArchiveストレージクラスが、選択される。

【 0 1 2 3 】

また、自記憶装置に他の記憶装置を接続することができ、記憶装置毎の特徴の違いを利用した記憶階層を構築することができる。そして、マイグレーションプランにそって、自記憶装置内ばかりでなく、他の記憶装置内のLUへファイルを移行することができ、1台の記憶装置で階層記憶制御を実現する場合と比較して、一層記憶装置にかかるコストを最適化することができる。

【 0 1 2 4 】

また、ストレージクラスが「Archive Storage」であるLUを構成するディスク装置は、ドライブを停止させてしまい、低消費電力を実現するとともに、ディスク寿命を延ばすようにすることも可能である。

【0 1 2 5】

また、さらに低コストな記憶装置を、本実施形態の記憶装置STR1に接続することともできるため、複数の記憶装置間により一層深い記憶階層を設定し、階層記憶制御を実行することでさらにコストを最適化することができる。

【0 1 2 6】**[実施形態 3]****(1) システム構成の一例 (図 1 4)**

次に第 3 の実施形態におけるシステム構成の一例を図 1 4 を用いて説明する。本実施形態においては、第 2 実施形態と同様、記憶装置STR0(1)に別の記憶装置STR2(1b)がネットワークを介して接続されたシステムにおいて、記憶装置間での階層記憶制御が実行される。第 2 の実施形態との相違は、記憶装置間を接続するネットワークが、第 2 の実施形態ではLAN20であり、記憶装置間ではファイルI/Oインタフェースが用いられるのに対し、第 3 の実施形態では記憶装置間を接続するネットワークが記憶装置間の接続専用のSAN35であり、記憶装置間ではブロックI/Oインタフェースが用いられることである。

【0 1 2 7】

図 1 4 において、記憶装置STR2(1b)は、第 2 の実施形態の記憶装置STR1(1a)と同様の小規模な構成の記憶装置であるが、第二の実施形態における記憶装置STR1(1a)のNASコントローラNCTL0に代えて、SANコントローラFCTLxを有している。FCTLxはSAN35に接続できるよう、FCコントローラ11012bを備えており、また、STR1のようなファイルアクセス制御用CPU11001aやその周辺回路が存在せず、ファイル制御は行わない。それ以外は、本実施形態における記憶装置STR2(1b)は、第二の実施形態における記憶装置STR1(1a)と同様の構成を有する。

【0 1 2 8】

SAN35は記憶装置STR0(1)と記憶装置STR2(1b)を接続するための専用ネットワークであり、SAN35にはSANホストは接続されていない。なお、本実施形態では説明の簡単化のため、記憶装置間を接続するネットワークであるSAN35にはSANホストが接続されないものとし、また、記憶装置間を接続するネットワークの本数も 1 本であるとしている。しかし、SAN35にSANホストを接続させることも、また、記

憶装置間を接続するネットワークの数を複数にして耐障害性を向上させることも可能である。

【0 1 2 9】

本実施形態においては、記憶装置STR0(1)の制御下に記憶装置STR2(1b)がおかれ、NASホスト0からのファイルアクセスは記憶装置STR0(1)を経由して記憶装置STR2(1b)に到達する。このような構成を以下「異種記憶装置接続」と呼ぶことにする。

(2)他の記憶装置へのファイルの移行処理

次にSTR0に格納されているファイルをSTR2移行するための処理について、第2実施形態との相違点を中心に説明する。

【0 1 3 0】

記憶装置STR0(1)のCHF1(1111)は、SAN35に異種記憶装置である記憶装置STR2(1b)が接続していることを認識する。CHF1(1111)はイニシエータとなり情報を採取するコマンドを発行することにより、STR2(1b)がSAN35に接続されていることを認識する。CHF1(1111)はSTR2の記憶領域をあたかも第1実施形態の自装置内部のディスクプールであるかのように扱う。CHN0(1110)はCHF1(1111)経由でこのディスクプールを使用することができる。ディスクプールの管理方法については後述する。記憶装置STR0(1)のCHN0(1100)はSTR2の構成を把握するため、自身がイニシエータとなり情報を採取するためのコマンドをCHF1(1111)を介してSTR2に発行する。そして、STR0のCHN0(1100)は、コマンドに対するSTR2からの応答をCHF1(1111)を介して受信し、応答に含まれる情報から、STR2がSATAディスクプールを所持し、15D+1P、RAID5構成で、2100GBの大容量を有する低コストなBlock TypeのLUを所持していることを認識し、このLUをリモートLUとして管理することを決定する。またSTR2が有するディスクプールは、大容量低コストなディスクプールなので、そのディスクプールのストレージクラスは「Archive Storage」とすることを、STR0のCHN0(1100)が決定する。そしてSTR0のCHN0(1100)は、STR2内のLUにLU4の番号を割り当て、このLUに関する情報、即ちStorage Class#「Archive Storage」、Storage Node#「STR2」、Disk Pool#「SATA Pool」、LU#「LU4」、LU Type「Remote Block」、RAID Conf.「RAID5 15D+P」、Usable Capacity「210

0GB」を、SM13のディスクプール管理テーブル131に格納する。STR0のCHN(1100)でファイルシステムプログラムが実行された場合には、ディスクプール管理テーブル131が参照され、このLUに関する情報がディスクプール管理テーブル131からファイルアクセス制御用メモリのストレージクラス管理テーブル1100451に複写される。

【0 1 3 1】

第2の実施形態と同様に、STR0のCHN0(1100)のマイグレーション管理部110043Aがファイルabc.docをNearLineストレージクラスからArchiveストレージクラスに移行することを決定したとして、以下にこの場合に実行されるファイルの移行処理を説明する。

【0 1 3 2】

STR0のマイグレーション管理部110043Aは、ストレージクラス管理テーブル1100439を参照し、「Archive Storage」クラスに該当するLU4を選択し、ファイルabc.docをLU4に移動することを決定する。LU4は、ストレージノードが「STR2(すなわち他の記憶装置1b)」、DiskPool#が「SATAディスクプール」、LU Typeが「Remote Block」という属性を備える。

【0 1 3 3】

第2実施形態とは異なり、LU4のLU Typeは「Block」型であるので、LU4にはファイルシステムが存在していない。そこで、CHN0(1100)に格納されているファイルシステムプログラムはこのLU4にローカルファイルシステムLFS4を構築する。LU4が設定されているディスクプールはSTR0から見ると他の記憶装置STR2に存在するので、「リモート」のディスクプールであり、LU4もリモートLUであるが、LU4内に設定されるファイルシステムLFS4の制御はCHN0(1100)が行うので、ファイルシステムLFS4はローカルファイルシステムになる。

【0 1 3 4】

LFS4はローカルファイルシステムとして管理されることになり、またLFS4が構築されるLU4はブロックタイプの異種記憶装置内に存在するLUであるので、本実施形態においては、ファイル格納管理テーブルの取り扱いが第1および第2の実施形態とは異なる。すなわち、STR0のCHN0(1100)のファイル格納管理部1100433

は、リンク先ノード名には「STR2」、リンク先F S名には「L F S 4」、リンク先ファイル名にはSTR2-FILE00001を割り当て、これを、abc.docについてのファイル格納管理テーブルに設定する。尚、abc.docは既にFILE00001のファイル名でLU2に移行されているので、CHN0(1100)は、割り当てられたリンク先ノード名、リンク先FS名、リンク先ファイル名を、LFS2のファイルFILE00001についてのファイル格納管理テーブルに設定しても良い。また、LU4が実際に存在するSTR2は、上述の通りファイルアクセス制御を実行しないので、STR2内にはSTR2-FILE00001に関するファイル格納管理テーブルが作成されない。

【 0 1 3 5 】

CHN0(1100)のファイルシステムプログラム11004が実行されることによって行われる処理は、ファイルabc.docのリンク先ノード(即ちファイルabc.docの実データが格納されている記憶装置)がSTR2であると意識されて処理が実行されること以外は、ファイルのオープン処理、ライト処理、マイグレーション処理ともに、第1実施形態で述べたローカルファイルシステムに対する処理と同様である。

【 0 1 3 6 】

しかし、ファイルを自記憶装置STR0内に存在するLUへ移動する第一の実施形態とは異なり、本実施形態では他の記憶装置STR2内に存在するLUへファイルのデータを移動するため、ディスクへの入出力処理は第一の実施形態とは異なる。第1実施形態ではディスクへの入出力処理をSTR0のDKA12xが制御したが、本実施形態の構成では、STR0のCHF1(1111)が制御する。そこで、CHN0のディスクアレイ制御用メモリ11009には、CHF通信ドライバプログラム110096が格納されている。ディスクアレイ制御用CPU11008がCHF通信ドライバプログラムを実行することにより、CHF通信ドライバ部が実現される。CHF通信ドライバ部は、SM13にディスク入出力コマンド(以下I/Oコマンドとする)を送信する。このI/Oコマンドには、データの格納位置を示すアドレス情報が含まれる。CHF1(1111)はこのI/OコマンドをSM13経由で受信し、受信したI/Oコマンドに基づいてSAN35経由で記憶装置1b(STR2)にI/Oコマンドを発行する。CHF1(1111)が発行するI/Oコマンドには、記憶装置1b(STR2)内でのデータの格納位置を示すアドレス情報が含まれている。記憶装置1b(STR2)は、通常のホストからディスクI/Oコマンドを受信した場合と同様の手順

で、CHF1(1111)から受信したI/Oコマンドを処理する。すなわちSTR2から見るとSTR0のCHF1がホストとして認識されることになる。

【0 1 3 7】

本実施形態によれば、ブロックI/Oインタフェースを備える異種記憶装置STR2が有するディスクプールを、記憶装置STR0のディスクプールの一つとして扱い、STR2が有するディスクプールに存在するLUに、STR0が管理するファイルシステムを構築することができる。更に、STR0内のLUに格納されているファイルをSTR2内のLUに移行することができ、コスト効果に優れ、柔軟な記憶階層を構築することができる。

〔実施形態 4〕

(1) システム構成の一例 (図 1 5)

次に第 4 の実施形態を説明する。本実施形態は、記憶装置の構成が前述の実施形態と異なる。

【0 1 3 8】

図 1 5 は、本実施形態におけるシステム構成の一例を示す図である。記憶装置STR3(1c)は、DKC70とディスクプールを備えている。DKC70において、SW71はスイッチ、NNODE (72x) はLANに接続するためのファイルI/O制御機構を備えたNAS ノード、FNODE(73x)はSANに接続するためのブロックI/O制御機構を備えたFC ノード、INODE(74x)はIPネットワークに接続するためのIPネットワーク制御機構を備えたIP ノード、DNODE(75x)はディスクプールを接続するためのディスク制御機構を備えたDisk Control ノードである。スイッチSW71には、一又は複数のNNODE 72x、一又は複数のFNODE73x、一又は複数のINODE74x、一又は複数のDNODE75xが接続されている。なお、スイッチにiSCSIを制御するノードを接続させ、IP SANを構築することもできる。iSCSIを制御するノードはFNODEと同様の機能を有し、同様の構成を有する。

【0 1 3 9】

DNODE0及びDNODE1は、FCディスクプール170とSATAディスクプール171の 2 種類のディスクプール 0 及びディスクプール 1 を接続し、制御している。

【0 1 4 0】

INODE0及びINODE1は、記憶装置STR3の外部に存在する、第2の実施形態で説明したファイルI/Oインタフェースを備えた記憶装置である、NAS型の異種記憶装置STR1(1a)と接続する。FNODE2及びFNODE3は、記憶装置STR3の外部に存在する、第3実施形態で説明したブロックI/Oインタフェースを備えた記憶装置である、SAN型の異種記憶装置STR2(1b)と接続する。

【0 1 4 1】

(2)NNODEの構成の一例(図16)

図16はNNODEの構成の一例を示す図である。NNODE720は図4に示すCHN1100のCPU間通信回路11007以下の構成要素を取り外し、SWノードコントローラ7404に代えたものと等価である。その他の構成要素については、CHNと構成、機能ともに同様である。

【0 1 4 2】

SWノードコントローラ7204はSW71と接続するためのコントローラ回路であり、コマンドやデータ、制御情報を、記憶装置STR3(1c)内部で送受信される内部フレームのフォーマットに形成し、DNODE等の他のノードにディスクI/Oとして送信する。

【0 1 4 3】

(3)FNODEの構成の一例(図17)

図17はFNODEの構成の一例を示す図である。FNODE730は図14のFCTL1100bのFCコントローラ11012bにSWノードコントローラ7302を接続した構成を有し、SWノードコントローラを介してSW71に接続できる。FCコントローラ7301はターゲットデバイスとして動作し、SANとの間でコマンドやデータや制御情報のフレームを授受する。SWノードコントローラ7302はFCコントローラが送受信するフレームを、記憶装置STR3(1c)の内部フレームの構成に変換し、DNODE等の他ノードとの間で送受信する。

【0 1 4 4】

また、FNODE73xはイニシエータデバイスとしても動作し、NNODEや他のFNODEから受信したディスクI/Oコマンドに基づいて、記憶装置STR3の外部に接続されている他の記憶装置にI/Oコマンドを送信することもできる。例えば、図15のFNO

DE2及びFNODE3は、記憶装置STR3のNNODEや他のFNODEから受信したコマンドに基づいて、記憶装置STR3の外部に接続された異種記憶装置STR2(1b)にI/Oコマンドを送信することができる。この場合、STR2から見るとFNODE2及びFNODE3はホスト計算機として動作しているよう見える。

【 0 1 4 5 】

図17では簡単化のため、FCコントローラ7301とSWノードコントローラ7302のみ記載しているが、ターゲット処理やイニシエータ処理、もしくは内部フレーム生成処理のために、CPUをFNODEに搭載する実装とすることも可能である。

【 0 1 4 6 】

尚、FCコントローラ7301に代えてiSCSIコントローラを設置すれば、iSCSIを制御するノードを構成することができ、このノードをSW71に接続すれば、IP SANを構成することもできる。

【 0 1 4 7 】

(4) INODEの構成の一例 (図18)

図 1 8 は INODE の構成の一例を示す図である。INODE740は図13に示すNCTL0(110 0a)が有するLANコントローラ7401にSWノードコントローラ7402を接続した行製を有し、SWノードコントローラを介してSW71に接続することができる。INODEは、STR3に外部NAS型の記憶装置STR1aを接続するために、記憶装置STR3(1c)に備えられている。

【 0 1 4 8 】

(5) DNODEの構成の一例 (図19)

図 1 9 は DNODE の構成の一例を示す図である。DNODE750は図 1 4 に示すFCTL110 0bのFCコントローラ11012bを取り外し、SWノードコントローラ7501に代えたものである。DNODE750は、ディスクI/OコマンドをSW71経由でNNODE又はFNODEから受信すると動作し、この結果、図 1 5 の点線で囲った1d部分が、あたかも独立した図 1 4 の記憶装置STR2のごとく動作する。本実施形態では、DNODE0(750)とDNODE1(751)とが二重化コントローラとしてペアを組み動作する。DNODEが二重化されている点も、FCTLが二重化されている図 1 4 の記憶装置STR2の構成と同様である。

。

(6) ファイルの移行処理

本実施形態は、第 1 の実施形態、第 2 の実施形態、第 3 の実施形態と記憶装置の構成が異なるだけで、階層記憶制御を実行するための処理手順は、第 1 の実施形態、第 2 の実施形態、第 3 の実施形態と同様であるため、記憶装置の構成の相違により生ずる動作上の相違点のみ説明する。

【 0 1 4 9 】

本実施形態においては、第 1 の実施形態と同様の手順で、記憶装置STR3の内部で階層記憶制御を実行することができる。NNODE72xのファイルアクセス制御用メモリに格納されているファイルシステムプログラム110043は、使用可能なLUを管理するためのストレージクラス管理テーブル1100439を備え、DNODE75xが管理するディスクプールおよびLUを、ストレージクラス管理テーブルを参照することで認識することができる。ただし、第 1 実施形態とは異なり共通情報を格納するSM 13が存在しないので、あらかじめNNODE72xが全DNODE75xに問い合わせ、使用可能なLUを特定し、ストレージクラス管理テーブルに登録しておく必要がある。もちろん本実施形態においても、SMを接続するためのSMノードを、SWに接続できるようにしておき、第 1 の実施形態と同様にSMに格納されている情報からストレージクラス管理テーブルを構成することとしてもよい。

【 0 1 5 0 】

NNODE72xが使用可能なディスクプールおよびLUを特定し、ストレージクラス管理テーブル1100439を作成し、ストレージクラスを定義すれば、後は第 1 実施形態と同様の処理にて、記憶装置STR3(1c)内での階層記憶制御、即ちディスクプール0及びディスクプール1内に設定されたLUを用いた階層記憶制御を実行することができる。

【 0 1 5 1 】

なお、ディスクI/Oコマンドの発行に関しては、NNODEのファイルアクセス制御用メモリ7203に格納されているSWノードドライバプログラムが、ファイルアクセス制御用CPU7202によって実行されることにより、SWノード経由でアクセス対象LUを管理するDNODE750に対しディスクI/Oコマンドが発行される。

【 0 1 5 2 】

以上のような構成、処理により、第 1 の実施形態の様な、記憶装置STR3の内部にファイルベースの記憶階層を構築したシステムを実現することができる。

【 0 1 5 3 】

また、第 2 の実施形態と同様に、記憶装置STR3の外部にファイルI/Oインタフェースを備えたNAS型の異種記憶装置STR1(1a)を接続し、記憶階層を構成することができる。NNODE72xのファイルアクセス制御用メモリ7203に格納されているファイルシステムプログラム110043は、ファイルアクセス制御用CPUによって実行されることにより、INODE74xに接続しているNAS型の異種記憶装置が存在するかどうかをINODE74xに問い合わせ、異種記憶装置が存在する場合には、この異種記憶装置内に存在するリモートLU、およびリモートファイルシステムを識別するための情報を異種記憶装置から取得する。そして、ファイルアクセス制御用CPUの制御によって、このリモートLUおよびリモートファイルシステムにストレージクラスを定義し、ストレージクラス管理テーブル1100439にこのLUに関する情報を登録して管理する。後は第 2 の実施形態における処理手順と同様である。

【 0 1 5 4 】

なお、ディスクI/Oコマンドの発行に関しては、NNODEのファイルアクセス制御用メモリ7203に格納されているSWノードドライバプログラムが、ファイルアクセス制御用CPU7202によって実行されることにより、NNODEからSWノード経由で、アクセス対象LUを備えた記憶装置STR1(1a)を接続するINODE740に対し、ディスクI/Oコマンドを発行する。INODE740は受信したI/Oコマンドに基づいて、ファイルアクセスを行うためのディスクI/Oコマンドを記憶装置STR1(1a)に発行するとともに、ファイルの実データや制御情報をSTR1(1a)との間で送受信する。

【 0 1 5 5 】

なお、INODE74xはファイルの制御情報には一切関与せず、単なるIPネットワーク用のゲートウェイとして動作する。この場合、他のNASホスト等のデバイスからの干渉を受けることがない階層記憶構成を実現できる。もちろん、第 2 実施形態と同様にNNODE720が接続するLAN20に異種記憶装置STR1(1a)を接続しても差し支えない。

【 0 1 5 6 】

以上のような構成、処理により第 2 の実施形態の様な、外部の異種記憶装置が有するストレージプールを用いたファイルベースの記憶階層を構築することができる。

【 0 1 5 7 】

また、第 3 実施形態と同様に、記憶装置STR3の外部にブロックI/Oインタフェースを備えた記憶装置である、SAN型の異種記憶装置STR2(1b)を接続し、記憶階層を構成することもできる。NNODE72xのファイルアクセス制御用メモリに格納されているファイルシステムプログラム110043が、ファイルアクセス制御用CPUによって実行されると、NNODEはFNODE73xに接続しているSAN型の異種記憶装置が存在するか否かをFNODE73xに問い合わせる。そして異種記憶装置が存在する場合には、NNODEは問合せに対するFNODEからの応答内容に基づいて、この異種記憶装置が有するリモートLUを認識し、このリモートLUにローカルファイルシステムを構築する。そしてNNODEは、このリモートLUとローカルファイルシステムにストレージクラスを定義し、ストレージクラス管理テーブル1100439にこのLUに関する情報を登録して管理する。後は第 3 の実施形態と同様の処理が実行される。

【 0 1 5 8 】

なお、ディスクI/Oコマンドの発行に関しては、ファイルアクセス制御用CPU7202がSWノードドライバプログラムを実行することにより、NNODEからSWノード経由で、アクセス対象LUを備えた記憶装置STR2(1b)を接続するFNODE732に対し、ディスクI/Oコマンドが発行される。FNODE732はディスクI/Oコマンドを記憶装置STR2に発行するとともに、データおよび制御情報を記憶装置STR2との間で送受信する。以上のような構成、処理手順により、第3実施形態と同様、外部の記憶装置STR2に構築され、記憶装置STR3によって管理されるファイルシステムを利用した、ファイルベースの記憶階層を構築することができる。

【 0 1 5 9 】

本実施形態によれば、記憶装置STR3があたかも階層記憶システムを構築するための中央制御コントローラであるかのように動作し、その記憶装置STR3の内部と外部に様々な種類の記憶装置を接続できるので、きわめて柔軟かつスケーラブルな大規模な階層記憶システムを構築することができる。また、記憶装置STR3のSW

の一ノードとして、記憶装置の内部にも外部にもディスクや他の記憶装置を接続できるので、高速なデータ転送を行うことができる。

(7)さらなる応用

以上、第1から第4の実施形態においては、ファイルのデータライフサイクルを基準としてファイルの階層的な移行処理を実行するファイルの移動方法及び記憶装置を説明してきたが、それ以外の基準でファイルを移動させることや、複数の基準を併用することももちろん可能である。データライフサイクル以外の基準としては、ファイルのアクセス特性やLUの使用容量などが考えられる。そして、ファイルのアクセス特性やLUの使用容量に基づくマイグレーションプランを設け、ファイルの移動を制御することができる。

【0160】

ファイルのアクセス特性に基づくマイグレーションプランの例としては、ファイルのアクセス頻度が一定水準を超えたら、一階層上位のストレージクラスにファイルを再移動するプランや、シーケンシャルアクセスに特化したストレージクラスを設け、あるファイルに対するシーケンシャルアクセス頻度が一定水準を超えたら、そのストレージクラスにファイル移動するプラン等が考えられる。

【0161】

LUの使用容量に基づくマイグレーションプランの例としては、LUの使用容量が一定水準を超えたら、現行ライフサイクルが変化していないファイルであっても、そのLUに格納されているアクセス頻度の低いファイルや、作成日からの経過時間が長いファイルを一階層下のストレージクラスに移行するプランが考えられる。

【0162】

上述の実施形態におけるファイル特性管理テーブル1100438にはファイル毎の動的特性としてアクセス情報を管理している。またストレージクラス管理テーブルにはLU毎の総容量と使用容量を管理している。これらの情報を活用することで、上述のマイグレーションプランを容易に実現することができる。

【0163】

【発明の効果】

ホスト計算機に依存することなく、記憶装置内部での処理によって、ファイルが有する特性に応じた階層記憶制御を実現することができる。

【図面の簡単な説明】

【図 1】

本発明が適用される計算機システムの構成例を示す図である。

【図 2】

記憶装置の外観の一例を示す図である。

【図 3】

アダプタボードの外観の一例を示す図である。

【図 4】

N A S チャンネルアダプタの構成の一例を示す図である。

【図 5】

ファイルシステム制御用メモリに格納されているプログラムの一例を示す図である。

【図 6】

ディスクアレイ制御用メモリに格納されているプログラムの一例を示す図である。

【図 7】

ディスクプールと、LUと、ファイルシステムの、対応関係の一例を示す図である。

【図 8】

ストレージクラス管理テーブルの一例を示す図である。

【図 9】

ファイル名管理テーブルの一例を示す図である。

【図 1 0】

ファイル格納管理テーブルとバッファ管理テーブルの一例を示す図である。

【図 1 1】

ファイル特性情報管理テーブルの一例を示す図である。

【図 1 2】

ファイル格納管理テーブルの一例を示す図である。

【図 1 3】

本発明が適用されるシステムの第二の構成例を示す図である。

【図 1 4】

本発明が適用されるシステムの第三の構成例を示す図である。

【図 1 5】

本発明が適用されるシステムの第四の構成例を示す図である。

【図 1 6】

NASノードの構成例を示す図である。

【図 1 7】

ファイバチャネルノードの構成例を示す図である。

【図 1 8】

I P ノードの構成例を示す図である。

【図 1 9】

ディスクアレイノードの構成例を示す図である。

【符号の説明】

1…記憶装置

1100…N A S チャネルアダプタ

1110…ファイバチャネルアダプタ

120…ディスク制御アダプタ

13…共有メモリ

14…キャッシュメモリ

15…共有メモリコントローラ

16…キャッシュメモリコントローラ

170…ディスクプール、

400…N A S ホスト

500…S A N ホスト

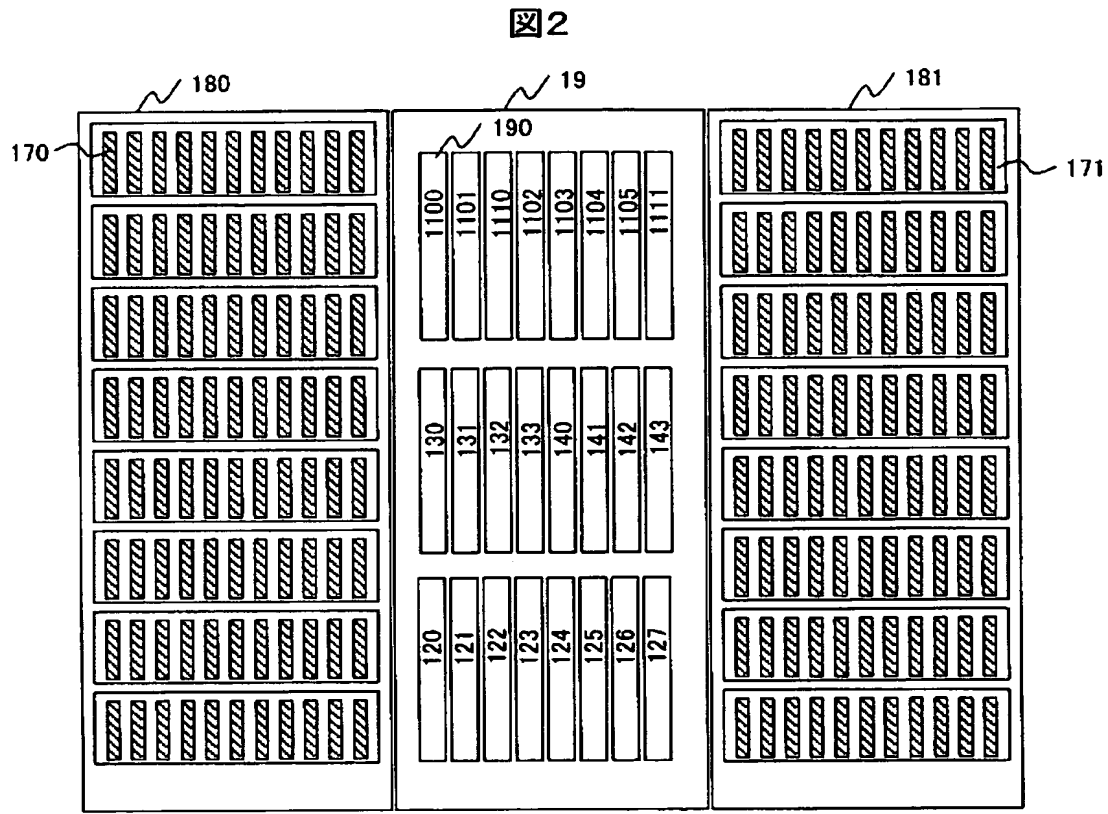
20…LAN

21…LAN

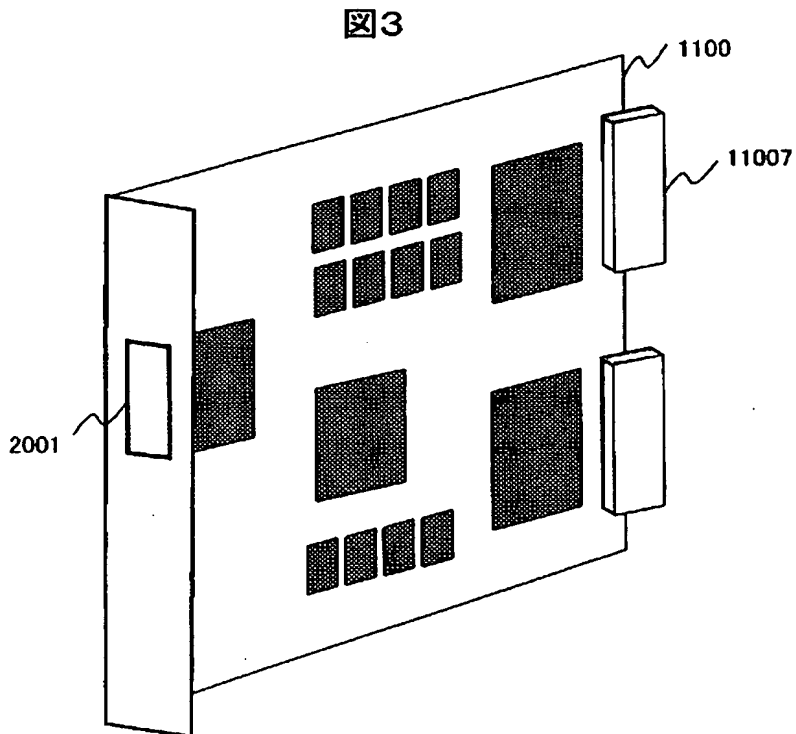
30...SAN

35...SAN

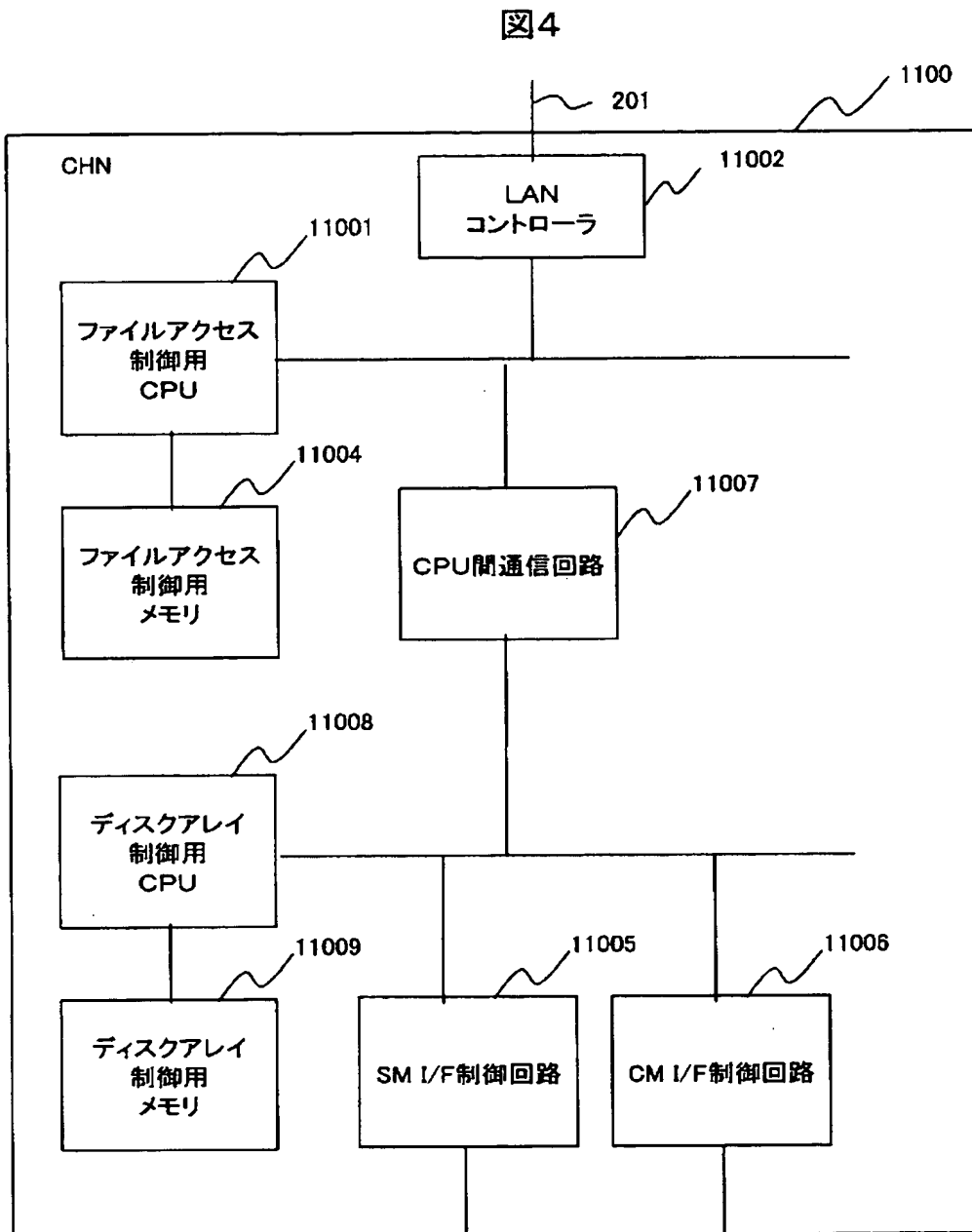
【図 2】



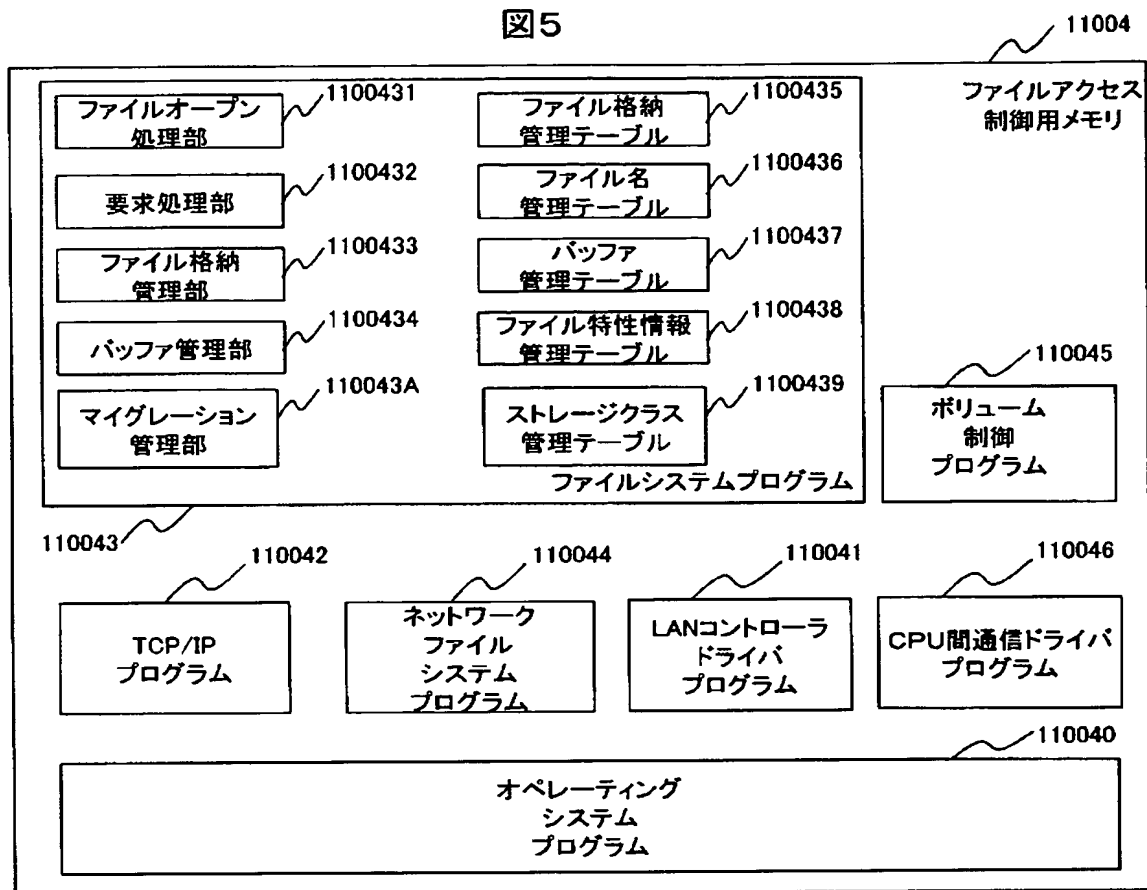
【図 3】



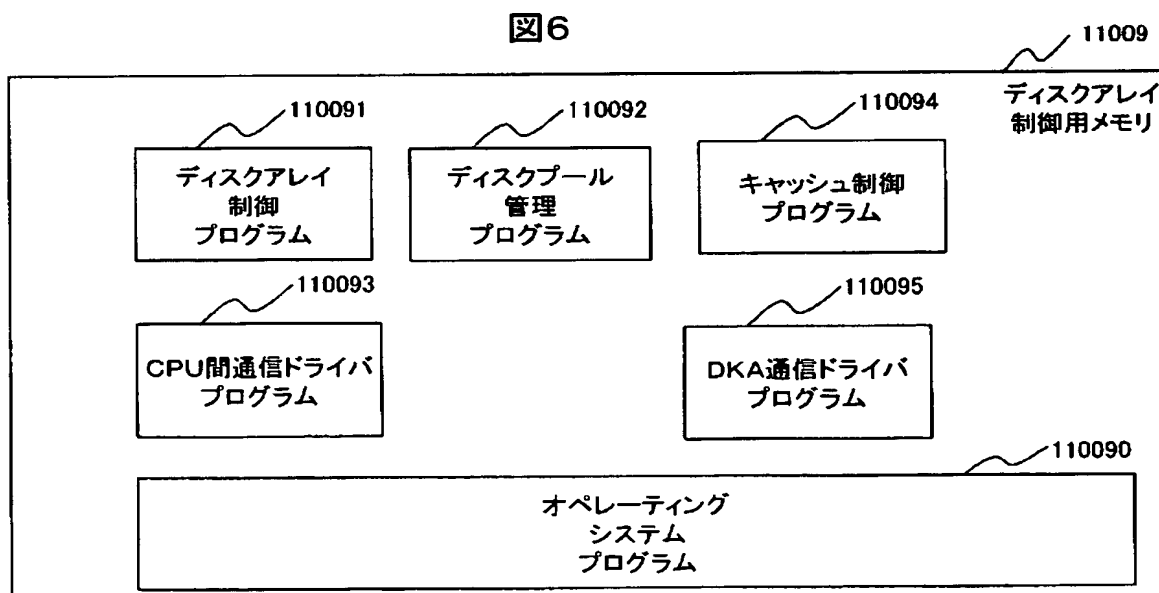
【図 4】



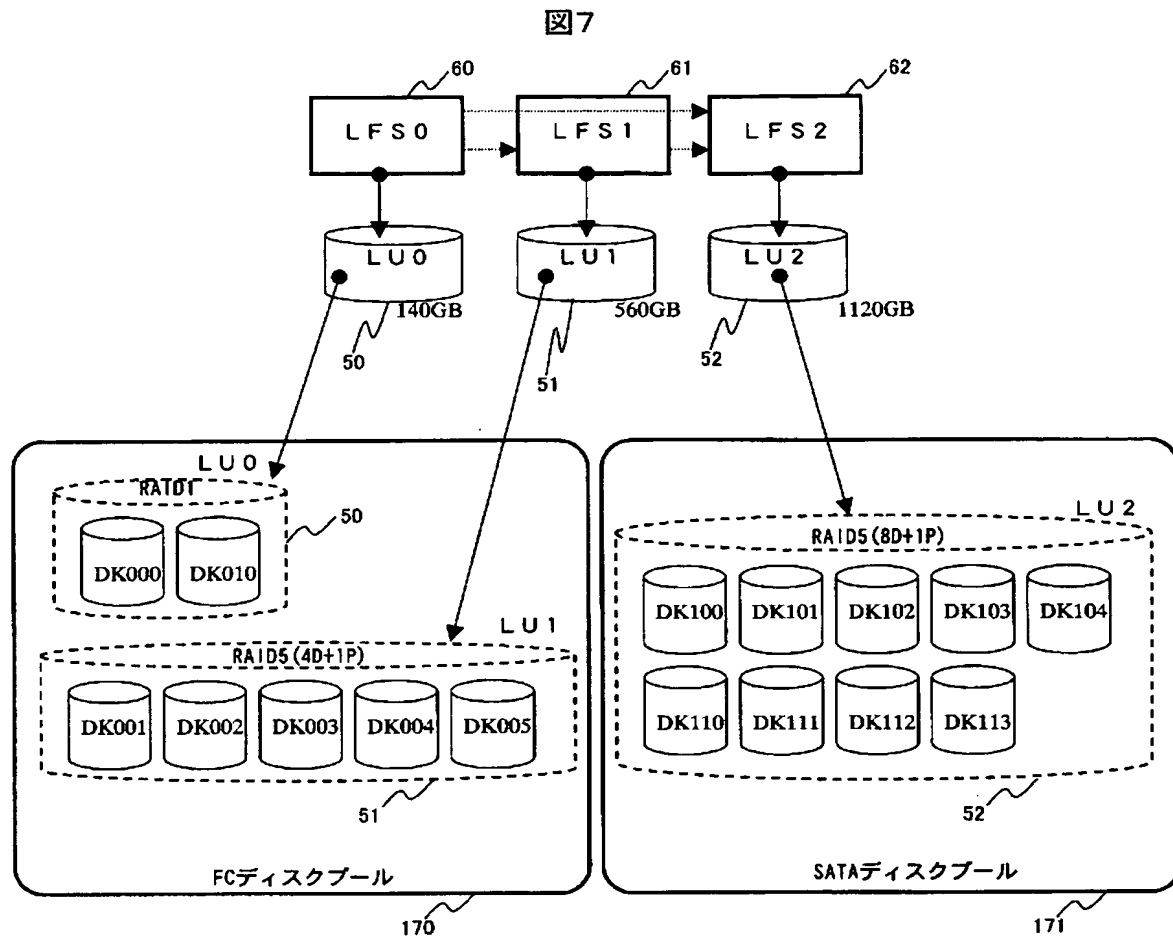
【図 5】



【図 6】



【図 7】



【図 8】

図 8

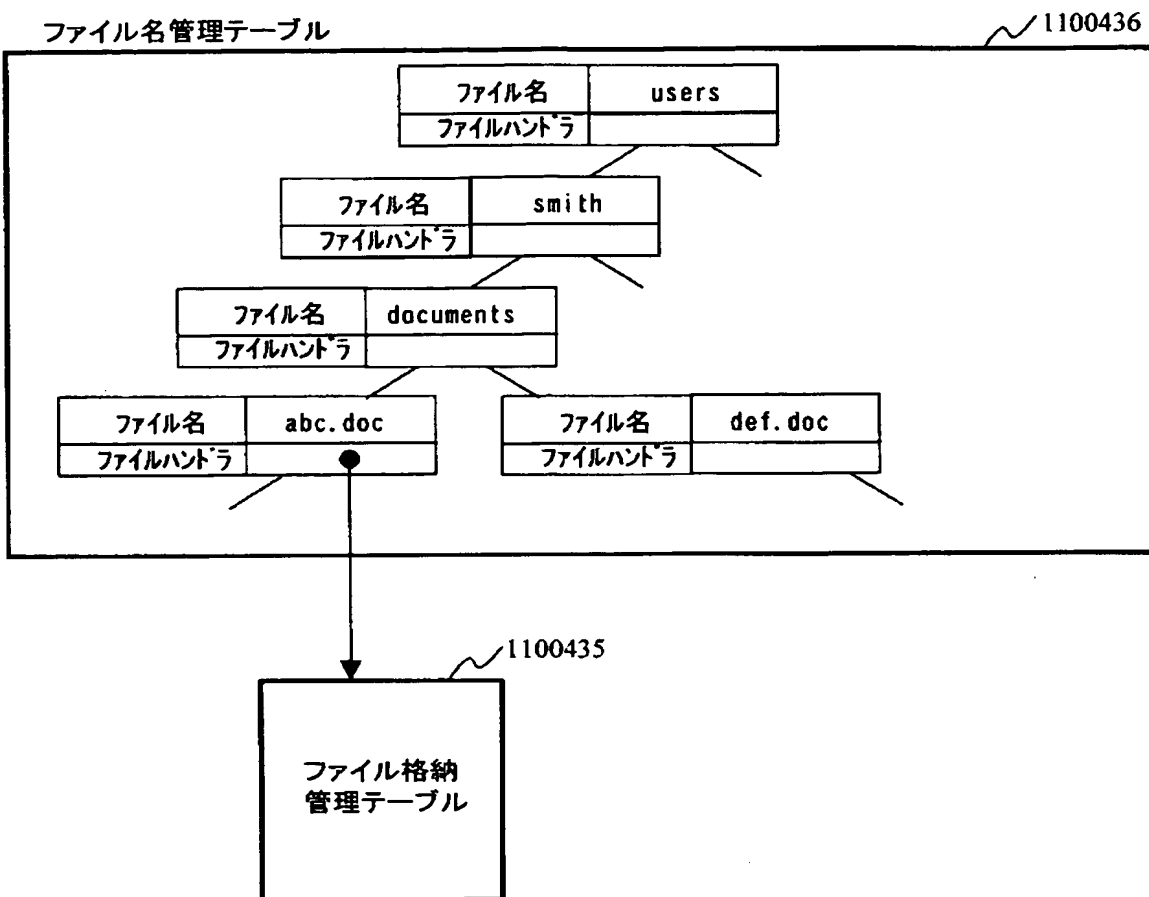
ストレージクラス管理テーブル

1100451

Storage Class# (1100451a)	Storage Node# (1100451b)	Disk Pool# (1100451c)	LU# (1100451d)	LU Type (1100451e)	RAID Conf. (1100451f)	Usable Capacity (1100451g)	Used Capacity (1100451h)
OnLine Storage (Premium)	STR0	FC Pool	LU0	Local File	RAID1 1D+1P	140GB	53GB
OnLine Storage (Normal)	STR0	FC Pool	LU1	Local File	RAID5 4D+1P	560GB	125GB
NearLine Storage	STR0	SATA Pool	LU2	Local File	RAID5 8D+1P	1120GB	320GB
Archive Storage	STR1	SATA Pool	LU3	Remote File	RAID5 15D+1P	2100GB	480GB
Archive Storage	STR2	SATA Pool	LU4	Remote Block	RAID5 15D+1P	2100GB	480GB

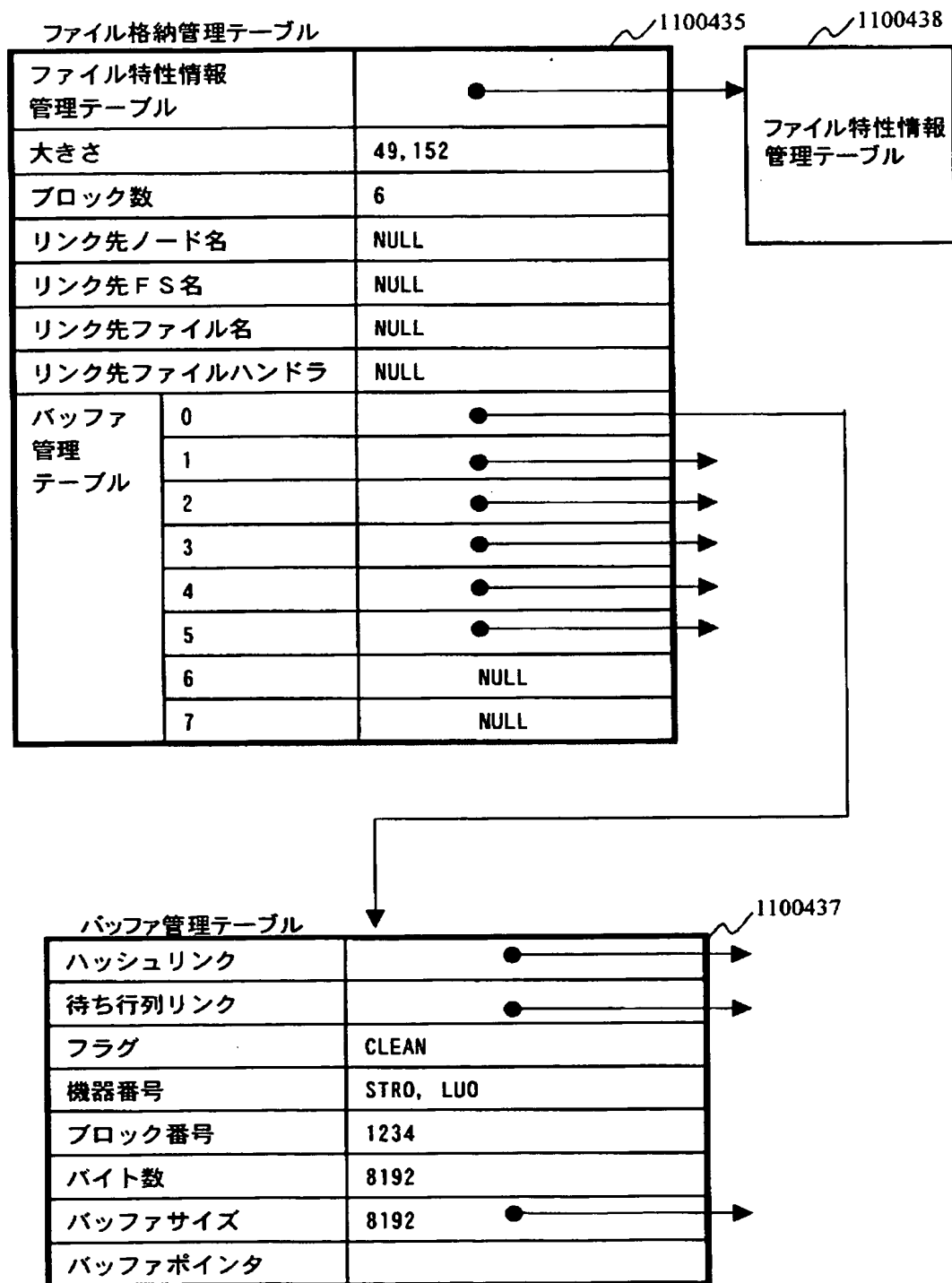
【図 9】

図9



【図10】

図10



【図 11】

図11

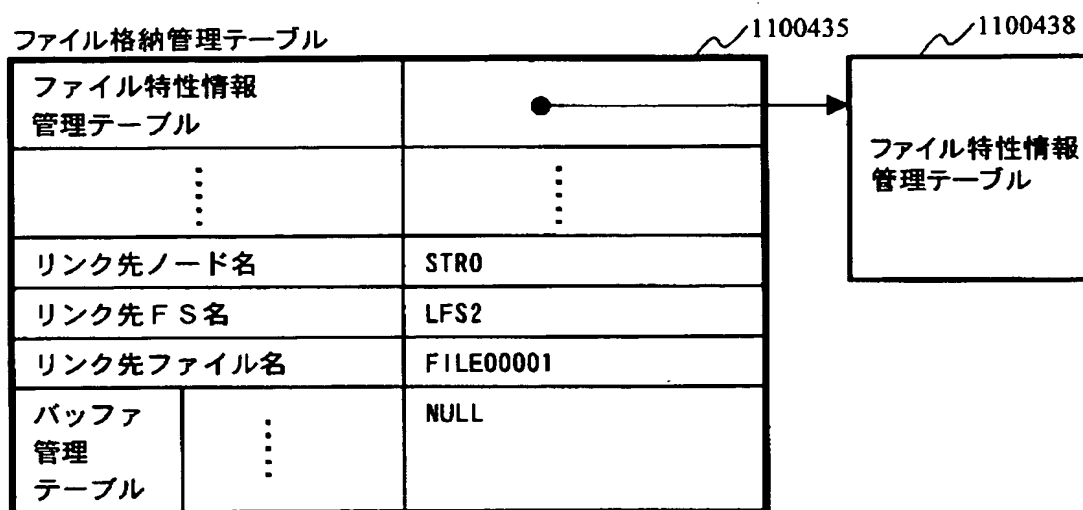
ファイル特性情報管理テーブル

1100438

特性情報種	カテゴリ	属性	内容
静的特性情報	ファイル情報	ファイル種類	ドキュメント
		アプリケーション	XYZワード
		作成日	2002.11.25
		所有者	Smith
		アクセス識別子	-rw-rw-rw-
	ポリシー	初期ストレージクラス	指定なし
		資産価値タイプ	重要
		ライフサイクルモデル	モデル1
		マイグレーションプラン	プラン1
動的特性情報	アクセス情報	タイムスタンプ	2003.1.1
		アクセスカウント	12
		リードカウント	7
		ライトカウント	5
		リードサイズ	7,168
		ライトサイズ	5,120
		リードシーケンシャルカウント	6
		ライトシーケンシャルカウント	4
	ライフサイクル 情報	現行ライフサイクル	参照期
		現行ストレージクラス	NearLine

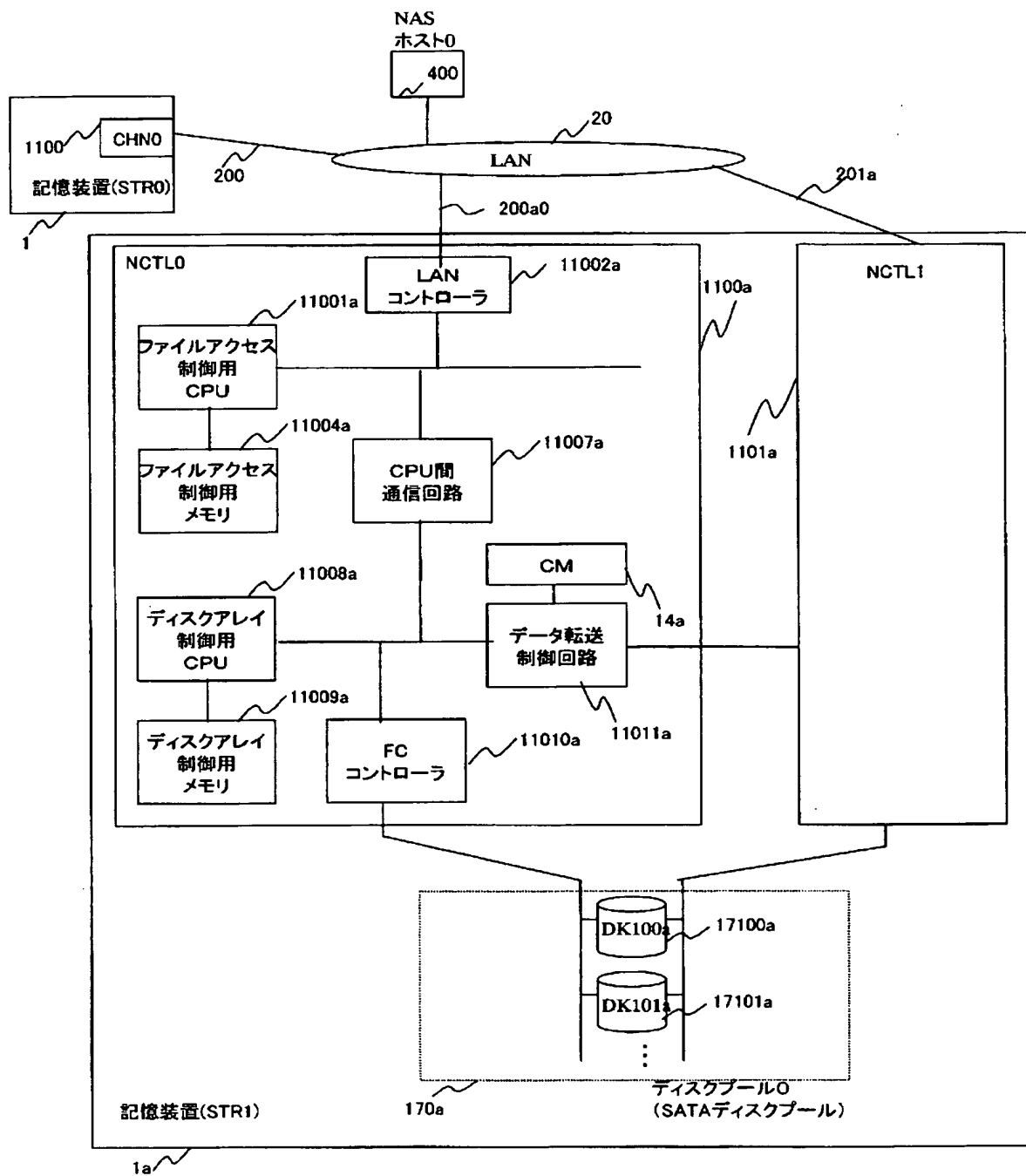
【図 12】

図12

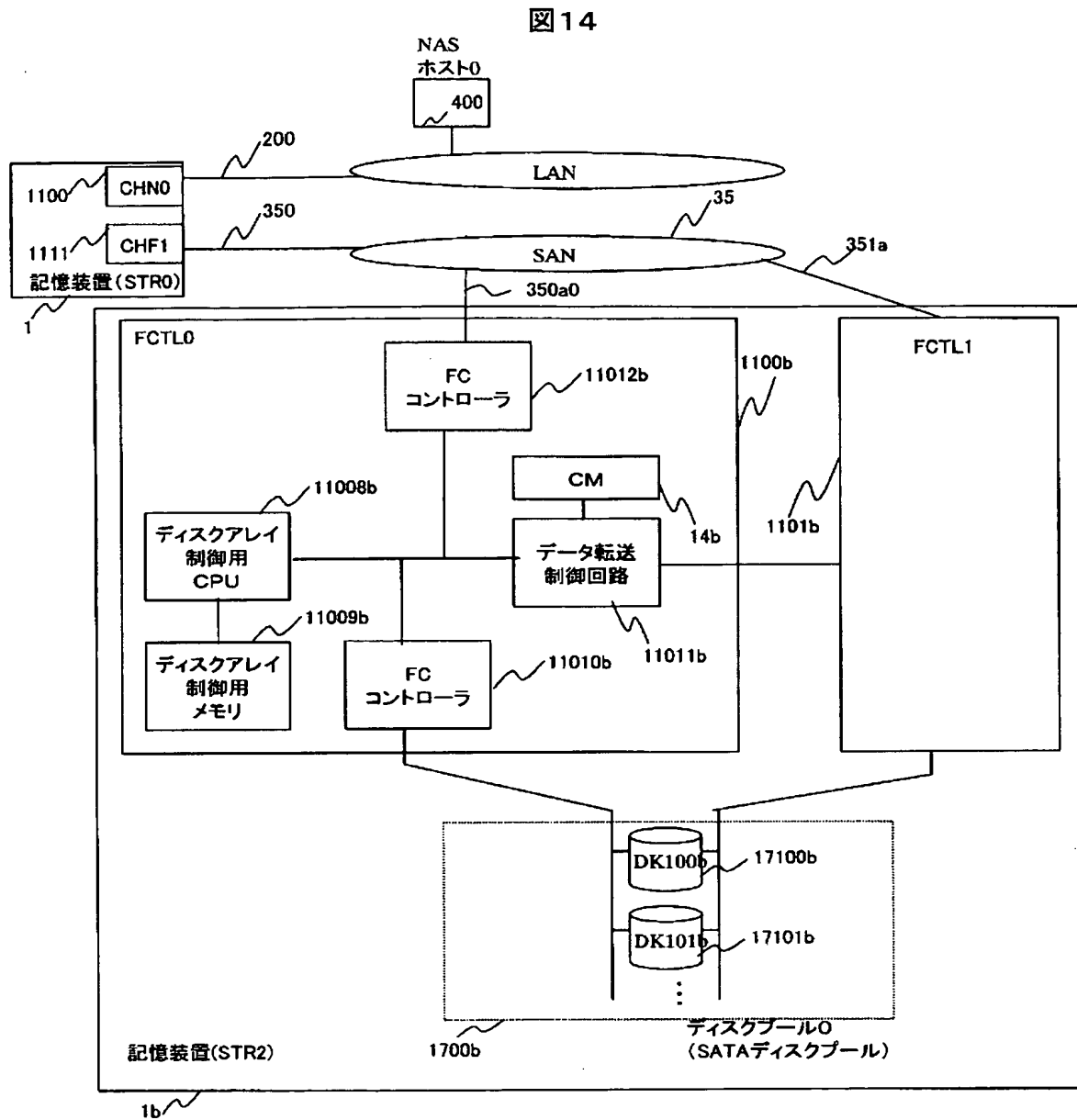


【図 13】

図13

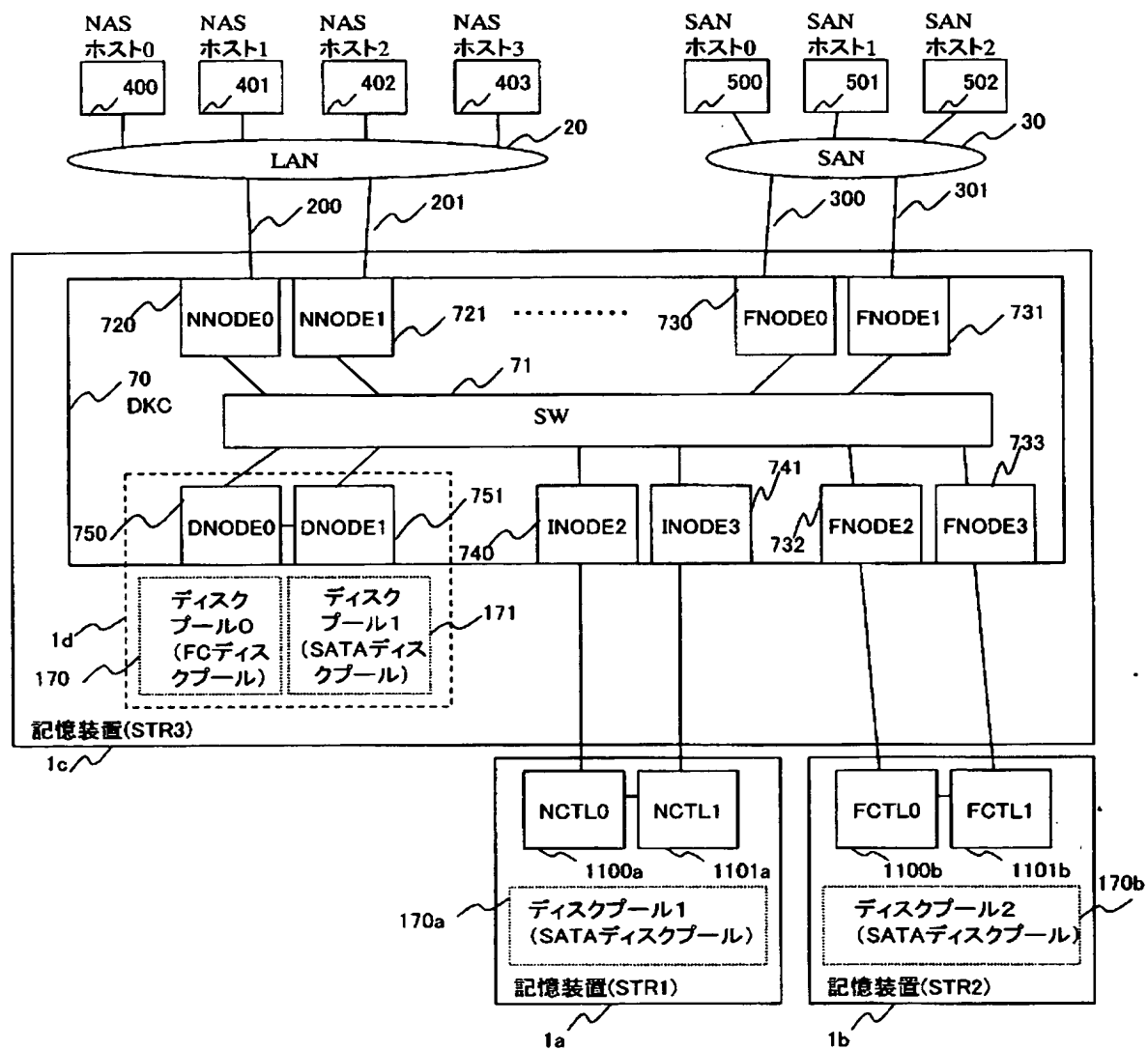


【図 14】



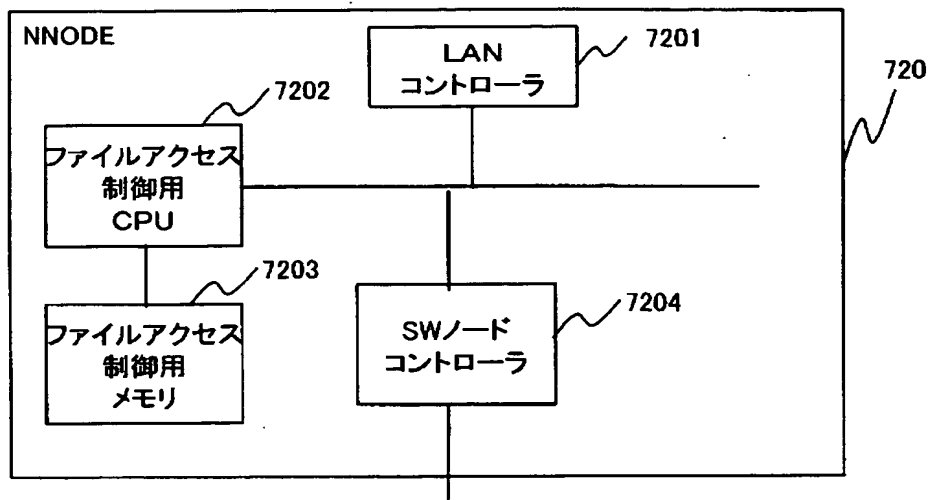
【図15】

図15



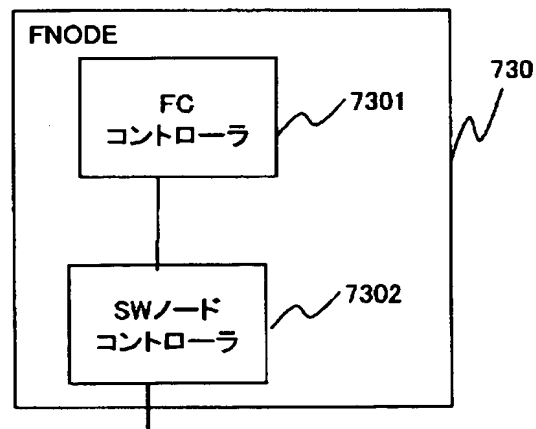
【図 16】

図16



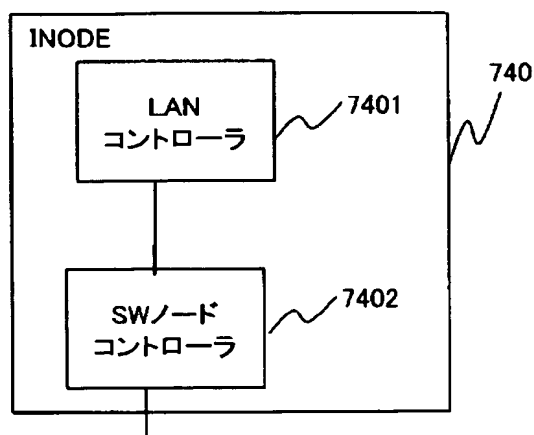
【図 17】

図17



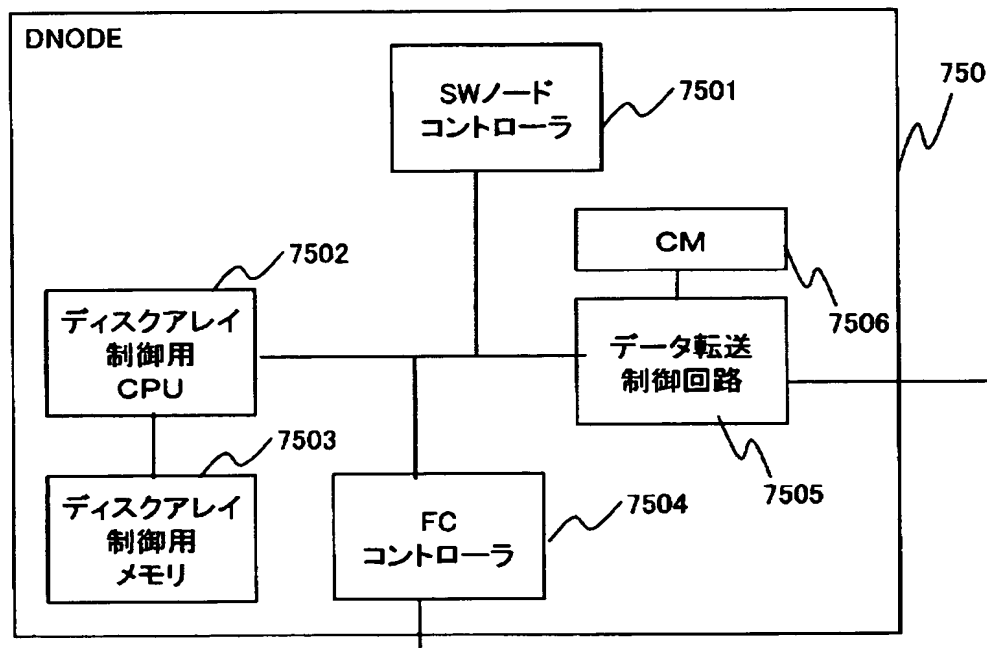
【図 18】

図18



【図 19】

図19



【書類名】 要約書

【要約】

【課題】

従来、記憶階層はホスト上プログラムにより実現されていたり、論理ディスク単位で実現されていたため、計算機への依存度が高く、柔軟なシステム構成や、管理の容易化に制限があった。また、ファイル特性に応じてファイルを最適配置することへの配慮がなかった。

【解決手段】

記憶装置は、ファイル I/O インタフェース制御装置と、複数のディスクプールを備える。ファイル I/O インタフェース制御装置は、ディスクプール内の LU に、複数のストレージクラスと呼ばれる記憶階層のいずれかを設定し、LU にファイルシステムを構成する。ファイル I/O インタフェース制御装置は、ファイルの静的特性とファイルの動的特性に基づき、最適なストレージクラスの LU にファイルをマイグレーションする。

【選択図】 図 7

認定・付加情報

特許出願の番号	特願 2 0 0 3 - 0 8 6 8 2 8
受付番号	5 0 3 0 0 4 9 9 3 5 0
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 5 年 3 月 2 8 日

< 認定情報・付加情報 >

【提出日】 平成15年 3月27日

次頁無

特願 2 0 0 3 - 0 8 6 8 2 8

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所